The National Hospital Peer Grouping Project

REPORTS
RA
981
A2
N38
1999

DEPARTMENT OF HEALTH & HUMAN SERVICES

Health Care Financing
Administration

Region VII
Federal Office Building
601 East 12th Street
Kansas City, Missouri 64106

Joe L. Tilghman
Regional Administrator
Health Care Financing Administration
Kansas City, Missouri

During the development of HCFA's Health Care Quality Improvement Program (HCQIP), a need for an objective and comprehensive statistical model for peer grouping acute care hospitals in the United States was realized. A workgroup was formed to develop such a model using existing HCFA administrative data and suitable data from the American Hospital Associations's surveys of hospitals. This work was contracted as a special project within HCQIP.

This document is the final report on the National Hospital Peer Grouping Project. Your attention is respectfully drawn to the workgroup's recommendations.

Kazim Sheikh, M.D.
Senior Medical Officer

Copy: Jeffrey L. Kang, M.D., M.P.H., OCSQ
Barbara S. Cooper, OSP

RA
981
.A2
N38
1999

The National Hospital Peer Grouping Project

Kazim Sheikh, M.D.
Health Care Financing Administration, Kansas City, Missouri

Kwan Lee, Ph.D., S.M.
William Rising, Ph.D.
Health Care Excel, the Peer Review Organization for Indiana and Kentucky

Deborah Felt, Ph.D.
Jane Deane Clark, Ph.D.
Michigan Peer Review Organization

March 1999

U.S. Department of Health and Human Services
Health Care Financing Administration
Baltimore, Maryland
March 1999

HCFA Pub. No. 10130

Project Workgroup

Jane Deane Clark received a Ph.D. degree in medical geography from Michigan State University; her dissertation was entitled *Variations in Michigan Hospital Use Rates: Do Physician and Hospital Characteristics Provide the Explanation?* Before joining the Michigan Peer Review Organization as the Executive Director of the Center for Health Outcomes and Evaluation, Dr. Clark was the Senior Director of Health Policy Analysis at the Michigan Hospital Association and Executive Director of the Michigan Health Data Corporation. She is currently the Vice President for Performance Measurement and Improvement at the Blue Care Network in Michigan, and she serves on the Boards of the American Nurses Credentialing Center and the Center for Clinical Quality Evaluation, American Hospital Association's Health Statistics Advisory Panel.

Deborah Felt received a Ph.D. degree in sociology with emphasis in research methods and statistics from the University of Minnesota; her dissertation was entitled *The Social Networks of Nursing Home Residents*. Before joining the Michigan Peer Review Organization as a Senior Staff Associate for Data Analysis, Dr. Felt held the position of a Quality and Marketing Statistician at the Memorial Hospital in South Bend, Indiana. She is currently a manager of health care projects at Orion Consulting, Inc. in Cleveland, Ohio.

Kwan Lee received a Ph.D. degree in biostatistics from the University of South Carolina; his dissertation was entitled *Small Area Rate Adjustments Employing Geographic Similarities and Empirical Bayes*. He also received a Master's degree in health policy and management from Harvard University with concentration in health economics and outcomes research. Dr. Lee worked as a biostatistician in the Medicare and Medicaid contracts of Health Care Excel, the Peer Review Organization for Indiana and Kentucky. He is currently an Associate Director of Statistical Analysis in the Joint Commission on Accreditation of Health Care Organizations' Department of Research and Evaluation.

William Rising received a bachelor degree in electrical engineering from the Massachusetts Institute of Technology and a Ph.D. degree in mathematics from the University of Massachusetts. Dr. Rising worked as a statistician at Health Care Excel, the Peer Review Organization for Indiana and Kentucky. He is currently a scientist in the Department of Biostatistics at Johns Hopkins University School of Public Hygiene and Public Health, Baltimore, Maryland.

The workgroup was assisted by Canopy Roychoudhury, Ph.D., Lisa R. Pressler, M.S., Heidi Zemke, M.S., and Ken Oishi, data analysts at the Michigan Peer Review Organization.

Acknowledgments

This project may not have been successful without access to the hospital survey data and a "crosswalk" to link these data with the Health Care Financing Administration's (HCFA) data, which were provided by the American Hospital Association (AHA).

Claudia Bullock at HCFA's Regional Office in Kansas City created customized files of HCFA's inpatient and physician identification data for this project, and she transported these data files and AHA's hospital survey data and computer program to the workgroup. Ms. Bullock's assistance is appreciated.

We thank the staff of HCFA and the Peer Review Organizations (PRO) for their very useful comments on the preliminary results of this project which were presented at HCFA's Grand Rounds and the American Health Quality Association's Technical Conference in 1997.

We are grateful to the PRO personnel who face-validated the hospital peer groups produced by the cluster analyses in this project:

- Gary Spivey, M.D., M.P.H., James Allen, M.D., and Barbara LaClair at the Kansas Foundation for Medical Care
- Velvet Yeager at Health Care Excel, Inc., Kentucky
- Ken Oishi at the Michigan Peer Review Organization
- Harriet Starr at the Island Peer Review Organization, Inc., New York
- Dale Bratzler, M.D., Lisa Bumpus, Kathryn Morris, Cynthia Murray, Scott Scheffler, and Jan Webb at the Oklahoma Foundation for Medical Quality
- Holmes Peacher-Ryan, Lesley Armishaw, and Maggie Blackhurst at the Mid-South Foundation for Medical Care, Tennessee
- Sharon Eloranta, Deena Uberaga, Sandi Salinas, and Jane Ballard at PRO-WEST, Washington

Executive Summary

1. The Health Care Financing Administration (HCFA) and the health insurance industry in the United States have categorized hospitals with similar characteristics into groups to compare the services provided and their cost. HCFA and several Peer Review Organizations (PROs) have also peer grouped acute care hospitals in HCFA's Health Care Quality Improvement Program (HCQIP). In most instances, hospitals in a state have been subjectively assigned to groups based on their size, urban or rural location, and/or their teaching or non-teaching status. Pitfalls in subjective peer grouping are well recognized. The importance of several other hospital characteristics in peer grouping is also well known.

2. An objectively developed peer grouping method which is simple, reliable, reproducible, adaptable to different situations, and which uses all important variables characterizing hospitals would have many uses. These would include appropriate comparisons of hospitals in HCQIP, marketing and evaluation of quality improvement interventions in hospitals, risk adjustment, and health services research.

3. For this project, four experienced staff statisticians at two PROs formed a workgroup to develop a statistical model for peer grouping acute care hospitals participating in the Medicare Program. HCFA's administrative data and data from the American Hospital Association's (AHA) 1994 survey of member hospitals were used in this project. These data were successfully merged for 5,205 acute care hospitals in the fee-for-service sector, but 4,444 hospitals for which complete information was available were included in the statistical analyses.

4. The development of the desired hospital peer groups was essentially a trial and error process. Two one-level and two two-level models were tried. In each model, a factor analysis of selected variables (hospital characteristics) followed by a cluster analysis were conducted. The factor analysis used the values of correlated variables to produce factor(s). The cluster analysis used the factor score(s) for each hospital and clustered them into peer groups. The face validity of the results from each model was examined, and the bootstrapping technique was used to objectively determine the proper number of peer groups using the structure of the data.

5. Data on 78 different and potentially useful variables were examined. Fifty-eight variables were found to be redundant because either their frequency distributions were skewed or they were highly correlated to other variables. Some variables, such as those representing gerontology services, were reduced to a single variable.

6. The first one-level model and the two two-level models were rejected because of their poor face validity, or because the resulting peer groups were predominantly determined by a single variable, or there were too many peer groups, or they were ill-defined. The maximum likelihood factor analysis followed by a cluster analysis were used in the last model. In this model, one

dominant factor explained 99 percent of the variation in all of the 20 variables entered in the analyses. The resulting 23 clusters (peer groups) were found to be well defined and valid.

7. However, the face validation showed that 23 peer groups were too many for states with small numbers of hospitals, because there were a few or no hospitals in some of the groups. These 23 clusters were subjected to another cluster analysis which produced seven large, valid clusters (peer groups). A comparison of the 23- and 7-cluster models suggested that the 23 clusters were in effect aggregated into seven large clusters. Thus the users have the choice of two (23- and 7-cluster) models.

8. A user's guide is appended to this report, and the users will be provided with a SAS computer program for updating or modifying their peer grouping models. All PROs will also be provided SAS transport files containing tables of hospitals in their specific states clustered into seven as well as 23 peer groups. These files will permit minor manipulations without re-running the cluster analysis.

9. Sixteen percent of the hospitals were excluded from the analyses because of missing information. These and other types of exclusions of hospitals may have created biased peer groups. HCFA's Provider of Service file was found to be an unreliable source of data because it was out-of-date, incomplete, and inaccurate. It is recommended that HCFA implements a process of improving the quality and completeness of its data sets so that future users of the peer grouping model, including HCFA, are not dependent on AHA data.

10. It may be that the AHA data will always be needed for updating hospital peer groups. Consequently, the workgroup recommends that HCFA and AHA improve the "crosswalk" for linking their data. By providing accurate and complete data to AHA and HCFA, hospitals can also facilitate improvement of the two data sets.

11. A major limitation of the data used in this project was its age. Many changes have occurred in the health care industry in the United States since 1994, and hospital peer groups based on 1994 data may not be valid today. This was apparent in the face validation process. It is recommended that the cluster analysis should be rerun using recent data, and that an ongoing process of updating the peer groups should be implemented.

12. The peer grouping methodology developed in this project would be very useful. HCFA, PROs, and hospitals should use these models in their policy development, health care and operational research, and in designing, marketing and evaluating health care quality improvement projects.

Table of Contents

Project Workgroup ..... *ii*

Acknowledgments ..... *iii*

Executive Summary ..... *iv*

Table of Contents ..... *vi*

1. Introduction 1

2. Plan 4

3. Background 5

4. Data 8

5. Methods 14

6. Results 19

7. Discussion 55

8. Recommendations 65

9. Glossary 68

10. References 69

Appendix A. Summary of Published Work on Peer Grouping Hospitals

Appendix B. Summary of Peer Grouping Methods Used by PROs

Appendix C. Data Dictionary

Appendix D. Instructions for Face Validation

Appendix E. Statistics of Key Variables in Each Cluster of Hospitals

Appendix F. User's Guide

1. Introduction

Comparisons and variations are essential features of health services research and evaluation. These comparisons or variations may be between patients, physicians, other care givers, medical care processes, outcomes, hospitals, other health care providers, health care utilization, patient satisfaction, or health care cost. Historically, hospitals with similar characteristics have been categorized into groups by the health insurance industry in order to compare the services provided and their cost. The Health Care Financing Administration (HCFA) has also classified hospitals participating in Medicare for the purpose of monitoring cost.

HCFA's relatively new Health Care Quality Improvement Program[1,2] (HCQIP) has used simple procedures for peer grouping acute care hospitals to compare a hospital's performance on specific indicators of the quality of medical care with that of its peers. These procedures have stratified hospitals with respect to two or three characteristics -- the hospital size, its urban or rural location, and the billings for selected procedures. The need for peer grouping hospital-specific data was also realized in the Medicare Quality Indicator System[3] (MQIS), a sister program for the development of health care quality indicators. Since many other factors, such as the number of services provided, case mix index, patient days, and the length of stay, can better discriminate hospitals, it was desirable to develop a statistical model which could make use of additional information on hospitals and objectively peer group them.

Ideally, there should be a standard method for peer grouping hospitals in any given situation. The method should use all available data on hospitals, the services they provide, the patients they serve, and the medical personnel who care for the patients. The peer grouping of hospitals should be based on the importance of the variables objectively selected by a statistical model. Yet, the model should be adaptable in different situations given different sets of variables -- for example, speciality hospitals, unusual clinical topics, selected geographical, political or administrative areas. Most importantly, the method should be simple enough and user-friendly for most data analysts to use. A reliable, versatile hospital peer grouping model would have the following uses:

1.1. Health Care Quality Improvement Program

1.1.1. HCFA's Medicare Quality of Care Surveillance System was initiated in 1995. It provides annual national and state-specific rates of health care utilization and outcome. These measures and their patterns and trends will be used by HCFA and other government agencies to develop policies and plans for targeting programs and for tracking progress. In order to interpret inter-state differences, peer grouping across states can identify comparable states in terms of the hospitals within the states. Researchers can use the inter-state comparisons to generate hypotheses for future research. The Peer Review Organizations (PROs) can also compare the measures for their states with those for their peers. If possible, PROs can compute the surveillance measures for groups of providers within their state and compare one group with others. This would allow them to identify and prioritize their quality improvement activities, as well as evaluate the statewide impact of selected HCQIP projects.

1.1.2. Before initiating local HCQIP projects, PROs often conduct preliminary studies to identify the problem areas, opportunities for improvement, and the providers to be targeted. Inter-hospital variations often indicate an opportunity for improvement. Peer grouping is essential for identifying comparable hospitals so that the observed variations among hospitals could be correctly interpreted.

1.1.3. Successful marketing is a critical step in conducting HCQIP projects. Traditionally, PROs have presented the baseline data on quality indicators in a project to potential collaborators in an attempt to demonstrate variations between the index hospital(s) and other hospitals in the state. These comparisons have not been acceptable when the comparability of hospitals was not convincing. To achieve voluntary collaboration of providers, PROs must demonstrate significant variation among comparable hospitals (the index hospital and its peers) and associated opportunity for improvement in health care. Consequently, PROs need convincing evidence by objectively peer grouping hospitals which can become a part of their protocol for local projects.

1.1.4. HCFA and the PROs who pilot national HCQIP projects also need peer grouping methodology to demonstrate opportunities for improvement and to achieve the collaboration of providers.

1.1.5. Hospitals within a peer group can compare their performance in terms of the quality of care with that of a benchmark institution both within and outside the peer group to determine the extent of possible improvement.

1.1.6 Since HCQIP projects are usually not designed as experimental studies or controlled trials, the effectiveness of the project interventions can not be objectively evaluated. Peer groups can be used to identify historical controls. Changes in health care processes attributed to interventions can then be compared with spontaneous changes in the control hospitals. It may also be possible to use peer grouping to match control state(s) with a state in which statewide interventions have been implemented. Consequently, hospital peer groups can be a very useful evaluation tool.

1.2. Medicare Quality Indicator System

An essential step in developing MQIS quality indicators is testing their reliability by comparing compliance with the indicators within a sample of hospitals. These samples are typically drawn from more than one state. The inter-hospital or inter-state comparisons can be meaningful only if the hospitals are comparable, that is, if the performance of a hospital or a group of hospitals is compared with that of its/their peers.

1.3. Risk Adjustment

Peer grouping of hospitals could be used to adjust risk in outcomes research. The purpose of adjustment would be to reduce confounding by external factors in comparison of more than one

set of data. Complex statistical methods are often required for risk adjustments. The statistical analyses can be simplified by selecting peer providers for comparison. The data on peers can also provide the standard from which the index hospital's expected range of performance can be estimated.

1.4. Research and Special Projects

Peer grouping of hospitals can be used to control confounding in the comparisons of outcomes and the prevalence of their risk factors associated with individual providers. For example, in case-control studies, peer groups could be used to match control hospitals with case hospitals. This procedure would have the potential for reducing the extent of necessary adjustments and standardizations in the analyses.

1.5. Reimbursement System

In addition to the location of a hospital, its peer group status could be factored into the reimbursement equation. This would likely improve the precision of a payment system and reward a hospital on the same scale as that for its peers regardless of their geographical location. To some extent, hospital peer groups are unintentionally included in the payment schedules. For example, the payment for teaching is the same regardless of the geographical location, and hospitals in the same Metropolitan Statistical Area are likely to be the members of the same peer group. Peer grouping of hospitals has the potential for being useful in developing payment schedules provided the model contains financial data and it is adequately tested.

1.6. Health Care Organizations Systems, Plans, and Associations

Hospital groups, health systems, hospital associations, and managed care plans need reliable methods for stratifying hospitals for the evaluation of services and cost. Peer grouping could also be used for continuous quality improvement activities.

2. Plan

A small group of experienced statisticians working in HCQIP at two PROs were selected and charged with using administrative data and appropriate statistical methods to develop a robust peer grouping model capable of stratifying acute care hospitals participating in Medicare into groups at the national and state levels. The model was to be so designed that it would primarily be used for evaluating and comparing the performance of hospitals in terms of the quality of health services they provide to Medicare beneficiaries. This developmental study was given the status of a special project within the current Scope of Work for those PROs. The members of the workgroup were expected to function as a team and accommodate this work in the course of their various routine work assignments.

After considerable deliberation, the following tasks were planned for the workgroup:

- search relevant literature;
- use personal knowledge and experience in hospital peer grouping methodology;
- improve on the existing peer grouping models;
- assess HCFA's existing administrative data for their suitability, and determine the need for and the sources of other easily available data;
- select and modify, if necessary, appropriate data management and statistical methods;
- interpret the findings and determine the need for further data manipulation;
- develop a scheme for and test the internal and external validity of the results;
- develop the criteria for and determine the pros and cons of the output of each model;
- decide on the final acceptable product;
- develop a users' guide and support service; and
- present and report the developmental process, its results, and specific recommendations to HCFA, PROs, and hospitals.

3. Background

The concept of hospital peer groups is commonly described in health services research literature. Furthermore, the inter-hospital comparisons ranging from utilization of services to quality of care or those related with hospital grouping methodologies have been frequently studied. In these studies, peer grouping of hospitals has been used to adjust for hospital characteristics that could affect the outcomes of interest so that statistical variations among the hospitals in a peer group would represent the true variations of interest. At the beginning of this project, the workgroup gathered information on peer group-related studies and experiences using various peer grouping methods which were previously published by PROs.

3.1. Literature Review

Extensive literature search identified 20 studies relevant to this project including six methodological studies. These are summarized in Appendix A.

3.1.1. Reimbursement System

Many previous peer groupings of hospitals were intended to estimate hospital costs so that each peer group could be used as a basis for determining reimbursement level.[4-8] These studies were often related to the external environments. They used various hospital financial and utilization data as the peer grouping variables which reflected the health care provided by the hospitals.[9, 10]

3.1.2. Inter-Hospital Comparisons

More recently, hospital peer groups have been used for inter-hospital comparisons of clinical outcomes or processes of care.[11, 12] In these studies, the expected range of hospital performances were estimated, for which the group average was often used as the norm. This provided an easy alternative to the model-based risk-adjustment that are relatively more resource-intensive.

3.1.3. Methodological Studies

There were two categories of hospital peer grouping methods in the literature -- simple stratifications based on a few hospital characteristics and multi-variate statistical methods utilizing multiple hospital characteristics at the same time.

3.1.3.1. Stratification Approach

Stratification has been a more frequent approach because of its simplicity. It is much easier to create hospital groups using this approach than the statistical methods, and when hospitals are stratified, the strata representing groups are usually easier to understand. However, this approach had to rely on a limited number of hospital variables, often not more than three. It was not realistic to characterize a complicated hospital structure with a few variables. In addition, when the variable of interest is continuous, for example the number of hospital beds, the variable needs to be categorized. In this situation, it would be inevitable to create "gray zones" in which

hospitals differing in certain stratification variables by a small unit could be stratified into different peer groups. Furthermore, peer groups created by this approach are generally difficult to test for their validity and reliability.

3.1.3.2. Multi-variate Statistical Methods

Multi variate statistical methods have been used to overcome the limitations of the stratification approach.[4-8, 11-14] The hierarchical cluster analysis has been the most widely used method. It is a multi-variate method since it allows the use of several hospital variables at the same time, without requiring any dependant variables.[15] In this approach, hospital groups are created through an iterative merging process in which similar hospitals are merged into the same peer group until certain stopping conditions are met. The stopping conditions were defined by entropy,[5] expected distinctiveness,[6] or R^2 statistic.[7] When the number of hospital variables is large, factor analysis was used to decrease the dimensionality before applying the cluster analysis.[6, 11, 14] When all of the hospital variables are binary, transformed values, such as Jaccard coefficients, can be used in the cluster analysis.[16]

3.1.3.3. Evaluating Hospital Peer Groups

Some studies have evaluated the integrity of peer groups after creating the groups. The evaluation methods were the analysis of variance[4, 7, 14] or the discriminant analysis[4] using hospital variables not used for the peer grouping. However, these approaches have limited applicability in evaluating reliability or external validity of peer groups. None of the published studies have utilized resampling methods such as bootstrapping, jackknife, or cross-validations.

3.1.4. Hospital variables used in peer-grouping

Various types of hospital variables have been used in previous hospital peer grouping studies. The most widely used hospital variables were the bed size, teaching status, and urban/rural location indicators. In general, hospital variables may be categorized as follows.

3.1.4.1. The hospital size has been measured by the number of hospital beds, the number of discharges, patient days, adjusted patient days, total full-time equivalent (FTE) employees, the number of FTE registered nurses, and total revenue.

3.1.4.2. Service availability or complexity is measured by the number of specialty services, availability of high-tech services, and the case mix index.

3.1.4.3. Hospital efficiency is measured by the cost per patient day, discharges per patient day, and the average length of stay.

3.1.4.4. Teaching hospital status is indicated by the ratio of interns/residents to hospital beds, and the ratio of interns/residents to average daily census.

3.1.4.5. Organization Structure is categorized as profit vs. not-for-profit, and private/public.

3.1.4.6. Geographic location is indicated by the population density of the area, HMO penetration, average income per resident, and the demographics of the residents in the area.

3.2. PRO/HCFA Peer Grouping Survey

During the Fourth Scope of Work, PROs were directed by HCFA to create state-specific hospital peer groups for use in HCQIP activities. PROs have developed several peer grouping models which were used as important analytic and reporting tools in HCQIP. Peer grouping has been used to identify outlier hospitals, to disseminate baseline data to state hospitals, or to assess hospitals' suitability for participation in HCQIP.

The peer grouping method most widely used by PROs was the simple stratification in which some combinations of the number of hospital beds, teaching status, and urban vs. rural status were used as grouping variables. HCFA's two national peer grouping schemes for its Cooperative Cardiovascular Project were also based on the stratification approach. In one scheme, HCFA created four peer groups using two peer grouping variables -- the number of beds and the urban indicator. In an other scheme, HCFA created three peer groups based on the number of invasive cardiovascular procedures.[17]

Some PROs used multi-variate statistical methods to create peer groups. The PRO for Indiana has developed a two-stage peer grouping method that combines the factor and the cluster analysis.[11] At the first stage, two factors were created based on 10 hospital variables; one reflecting hospital size and the other reflecting the complexity of services. Next, cluster analysis was conducted based on the factor scores, and it produced six peer groups based on 110 statewide hospitals. The Michigan PRO enhanced the multi-variate statistical methods with the development of a scaling system of specialty services.[12] A summary of PROs' peer groups is given in Appendix A.

4. Data

4.1. Data Sources

There were two sources of data for the peer grouping model -- the Health Care Financing Administration (HCFA) and the American Hospital Association (AHA). This section of the report will describe the data sets themselves, the manipulations used to create complex variables of interest, and the strategies used to link the data from the various files.

4.1.1. HCFA Data

4.1.1.1. The Inpatient Standard Analytical File contains administrative data from every bill submitted to HCFA for every hospitalization by every provider participating in the Medicare program for reimbursement as well as a relatively small number of no-pay bills submitted by Health Maintenance Organizations (HMOs). All records which arose from Prospective Payment System (PPS) stays, which were complete, valid, and for which the discharge occurred in calendar year 1994 were selected for this study. Only a subset of the fields available on each record were used. These fields contained demographic information, diagnosis and procedure codes, physician codes, and admission and discharge information. The Inpatient file was the primary file for information about the Medicare patients served by each hospital. This selection amounted to 10,995,623 discharges.

4.1.12. The Case-Mix Index (CMI) file provides a lookup table which gives the CMI for each Medicare provider for a government fiscal year. The CMI itself is the arithmetic mean of the Diagnosis Related Group (DRG) weight for the PPS discharges for the year. The DRG weight, in turn, is the multiplier by which the base reimbursement for a Medicare stay is multiplied. Thus, higher DRG weights correspond to more expensive diagnoses and procedures. The particular data set used for the peer grouping effort was for discharges occurring between October 1, 1993 and September 30, 1994. The CMI index was chosen to provide a measure of the intensity of care at each provider.

4.1.1.3. The Provider of Services (POS) file is HCFA's list of all Medicare providers and the services they provide. Because of earlier work which cast doubt upon the reliability of some of the POS fields,[18] and because AHA hospital data were available, this file was used only for linking data sets (See 4.3.4). The POS file was also from 1994.

4.1.1.4. The Physician Identification Master File for 1994 was acquired with the hope of getting information on 740,743 physicians working in the individual hospitals. This file contains a wealth of information about physicians participating in Medicare, including information about graduation year, board certifications, and years of practice. In the end, however, linking problems prevented use of the information in this file.

4.1.2. AHA Data

4.1.2.1. The AHA Annual Survey of Hospitals is a comprehensive database compiled yearly by the AHA from its member hospitals through a survey. The survey is completed by individual hospitals, checked for consistency by the state hospital association, and, if need be, revised to reflect accurate information. Approximately 90 percent of the member hospitals respond to the survey.[19] In case of item non-response, AHA estimated the missing values from previously reported data or through imputation techniques using related variables. The file contains information about hospital control and governance, its size, staffing levels, patient census, services provided, as well as some insight into the extent of outpatient services for each provider, expenses, hospital classification, and geographic codes. The strength of this database lies in its careful checking of errors, its weaknesses can be found in the refusal of some providers to answer some of the questions, and in its lack of complete coverage since not all hospitals belong to the AHA. In the 1994 survey, each of the 6,591 responding hospitals provided information on 594 variables.

4.1.2.2. The AHA Provider Crosswalk is a lookup table which links AHA provider numbers with HCFA provider numbers. This file was critical in linking the AHA and HCFA data together.

4.2. Data Additions and Manipulations

4.2.1. Patient Transfers

To capture a hospital's role as either a referral center or a tertiary care center, it was deemed necessary to look at the rate of transfers both out of and in to each facility. The source for this information was the inpatient data set. Transfers were marked in two fashions; the first based solely upon the information found on a single record, and the second exploiting the other records in the database. The two methods are necessary because of the limitations in the accuracy of administrative data. In the particular case of transfers, there is incentive for accurate coding of transfers going out, since HCFA disapproves discharges (non-transfers) which are closely followed by an admission. There is, however, no incentive for coding incoming transfers.

The first method defined transfers coming into a provider as any admission type other than a newborn admission (including missing values) together with an admission source of a transfer from a short-term general hospital. Outgoing transfers were defined as those discharges which had a discharge destination of another short-term health care facility.

The second method combined the above definitions of transfers with information about the previous and following discharges, if any. Transfers going out of the hospital were redefined to be those which had a discharge destination of a short term acute care facility as well as a following admission within two days of the discharge date. The new definition of a transfer coming into a hospital was to additionally mark those admissions whose previous stay had been an outgoing transfer as mentioned above. This method for marking the incoming transfers reflects the bias towards more accurate billing data for outgoing rather than incoming transfers.

4.2.2. Emergency Admissions

Emergency admissions were computed in two different fashions. Both of these may be questionable, since there is no incentive for correct coding, and both admission source and type are rather unreliable fields in the inpatient data set. The first step was to use the admission source of 7 (emergency) together with any admission type other than a newborn. The second was to use an admission type of emergency.

4.2.3. Inpatient Data at the Provider Level

Whereas the inpatient data arrived at the beneficiary level, the data needed for the peer-grouping effort were at the provider level. Thus, the data sets needed to be collapsed. The provider level information which was computed were various rates, mean length of stay, and DRG entropy. Table 1 lists the rates which were collapsed to the provider level. In addition to these rates, the arithmetic and geometric mean lengths of stay (LOS) were computed without adjustment for deaths.

Table 1. Rates collapsed to the provider level

Data	Numerator	Denominator
Females	Females	Discharges
Mortality	Deaths	Discharges
Emergency #1	Emergency admissions based on admission source	Valid admission source fields
Emergency #2	Emergency admissions based on admission type	Valid admission type fields
Caucasian	Caucasians	Discharges
Transfers in #1	See above	Discharges
Transfers out #1	See above	Discharges
Transfers in #2	See above	Discharges
Transfers out #2	See above	Discharges
Procedures	Discharges with at least one procedure code	Discharges

As the hospital market becomes more segmented, the need for measuring the level of specialization of hospitals becomes more critical for peer grouping. The measure chosen for this is the DRG entropy.

Entropy is a quantity from thermodynamics which measures the disorganization of a system. In more pedestrian terms for example, windy days increase hair entropy. This concept can easily be brought over to discrete probability distributions by defining the entropy of the distribution as $S = -\sum_i p_i \ln(p_i)$, where p_i is the probability of the i^{th} atom, and $0^0 \equiv 1$. Calculus shows that S is minimized at 0 when there is a single atom (whose probability is necessarily 1), and is maximized at $\ln(n)$ for n atoms each with probability $1/n$.

Applying this to the empirical or sampling distribution, the formula becomes

$$S = \ln(n) - \frac{1}{n}\sum_i n_i \ln(n_i)$$

where n is the total number of observations, and n_i is the number of observations in each category.

This quantity was computed for each hospital, using discharges as the counts, and DRGs as the categories.

4.3. Data Linking

4.3.1. Linking the AHA File to the Inpatient File

Linking data sets was clearly the central procedure needed to join the data from HCFA with the data from AHA. Ideally, this could be accomplished by using the AHA crosswalk file which links each AHA provider number with its corresponding Medicare provider number (MPN). In reality, however, there were providers which could not be linked by using the crosswalk, as well as providers which had non-unique entries. Various procedures used to increase the number of providers for whom the AHA data and HCFA data could be linked are described below.

4.3.1.1. Aggregating Facilities

The first quirk in the crosswalk file was one of multiple AHA hospital entities which corresponded to a single HCFA provider. This arose from multiple physical units which all billed as a single financial unit. In order to correct the problem, the names of each of the physical units from the AHA annual survey file were manually compared to the names of the financial units in the POS data. The correct 'flagship' unit was identified and flagged in each case. Subsequently, the data from the AHA annual survey had to be aggregated, so that the bed counts, available services, and the like would be correct for the single financial unit. The variables were aggregated according to their type. If the variable was logical (Yes/No), the results were combined with logical 'or's, so a single Yes for any unit yielded a Yes for the combination. If at least one response was missing and the rest of the responses were No, the

combined value was taken to be missing. Counts were combined by adding. Missing values were treated as zeroes. Two variables were neither simple counts nor logical. The bed size classification variable was revised by first computing the total number of beds, and then assigning the proper class.

4.3.1.2. Merging Missing Providers

In addition to the problem of missing MPNs, some duplicate MPNs were found. These providers appeared in the inpatient data but could not be linked using the crosswalk. There are three reasons for such gaps; (a) the providers may not be part of the AHA, so they would have no data in the annual survey, (b) the crosswalk could be missing the providers, though there are data in the AHA file, or (c) the providers may be in the crosswalk, but missing from the AHA file.

In an effort to merge as many providers as possible, a merge was performed using names and addresses from the AHA and POS file. First, the city, state, and the first six words in the names were matched, followed by the ZIP code and the first six words, followed by the state and the first six words. This trilogy of matching criteria was iterated down to just the first two words in the provider name. Lastly, matches on city, state, and ZIP code, and just ZIP code were made. In every case, a match was considered to occur if, and only if, there were single matches for the matching criteria (that is, ZIP and two words would be considered to match if, and only if, there was just one provider to match the particular combination of ZIP and two words), and if there were no more complex matches. A visual inspection of the matches was made to ensure that there were no egregious errors. This method matched 231 of the 410 hospitals which appeared in the inpatient data set but could not be matched using the crosswalk.

4.3.2. Linking the Case Mix Index to the Inpatient File

Linking these files was relatively easy, and all the providers matched without problems.

4.3.3. Linking the Physician File to the Inpatient File

The linking of the physician file to the inpatient file failed to be of any value, mostly due to problems inherent in the physician file. The variables which were problematic included the physician speciality certification code and the year of graduation. Forty-nine percent of speciality codes were unknown, and 12 percent of entries for the year of graduation were incorrect -- for example, graduating in 1995 or later or graduating in year zero. In addition to these basic errors, linking of the Unique Physician Identification Number (UPIN) and inpatient files failed due to a multitude of unmatched UPINs on the inpatient file. For these reasons, use of physician data from the UPIN file was abandoned.

4.4. The Final Data Set

The merges described above resulted in a data set with both inpatient and hospital information for most of the Medicare providers who submitted claims during the calendar year 1994. Of the

5,326 acute care hospitals submitting claims, only 121 were not matched in the AHA file. Of the remaining 5,205 hospitals, there were 761 whose missing data elements in the AHA file led to their exclusion in the actual clustering. Thus the final data set contained complete information on 4,444 Medicare providers.

5. Methods

The hospital peer grouping model was developed by (a) selecting the data variables which had significant independent effect on grouping of hospitals, (b) collapsing these variables into factors, (c) objectively determining the optimum number of peer groups, (d) clustering hospitals with similar characteristics into groups, and (e) validating the resultant peer groups.

5.1. Variables

HCFA's Medicare Inpatient Standard Analytic file was the source of the total number of Medicare patient discharges in calendar year 1994, and those discharges which were transferred out of the hospital (See Section 4.2.3). Claims that contained an ICD-9 surgical procedure code were used to estimate the percent Medicare discharges with a procedure code. DRG Entropy has been described above (Section 4.2.3).

The total number of inpatient days was estimated from the AHA's variable Inpatient Total Days. The adjusted patient days were defined as the sum of inpatient days and equivalent patient days attributed to outpatient services. The latter were derived by multiplying the inpatient days by the ratio of outpatient revenue to inpatient revenue. The AHA's 1994 survey data were the source of these measures.

The total number of set-up and staffed beds used as a variable had been calculated by AHA. The Average Daily Census was also calculated by AHA as the ratio of inpatient days to the number of days in the reporting period. The occupancy rate was derived by dividing the Average Daily Census by the bed total.

Eighteen services documented in the AHA file were combined to form one variable called the number of gerontology services. These services were radiation therapy, cardiac catheterization, angioplasty, cardiac intensive care, case management, oncology services, magnetic resonance imaging (MRI), open heart surgery, services for HIV-AIDS, single photon emission computed tomography, rehabilitation care, organ transplantation, extra-corporeal shock-wave lithotripter, freestanding out-patient center, geriatric services, general intensive care, certified trauma center, and computed tomography (CT) scan.

These variables were selected on the basis of their relevance to the gerontological population. For example, mother/baby services were excluded. The services that were most relevant to a gerontological scale were summed and evaluated using the Chronbachs alpha statistic which is available in SAS. It is a reliability coefficient and is used in psychometrics to construct reliable scales. The coefficient ranges between 0 and 1, the higher the fraction of 1 the better the scaling properties. A coefficient above 0.80 indicates that the variables can be scaled into one variable. The alpha coefficient for the gerontology services was 0.87.

Participation in HMOs and PPOs was defined as the hospital having a formal written contract with an HMO and PPO respectively that specifies the obligations of each party. The AHA

variables HMO86, IPHMOHOS, HMOCON, PPO86, IPPPOHOS, and PPOCON were used to construct these variable.

The AHA classification was used to categorize the MSA population size.

The AHA data were also the source of the membership in a health system, full-time-equivalent (FTE) registered nurses (RN), FTEs in a hospital, total medical residents, accreditation by the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO), and for-profit ownership of hospitals by individuals, a partnership, or a corporation.

5.2. Factor Analysis

The factor analysis is a data reduction technique that enables researchers to determine whether the measured variables can be explained by a smaller number of underlying "factors".[20] Each factor is a new variable that is correlated with the original variables. The correlation is in effect the factor loading, and it determines the many-to-one (common factor) or one-to-one (unique factor) relationship between the variables and the factor. Factor analysis uses all selected variables even though they may be highly correlated. The extracted factors are not correlated with each other. The number of factors is determined based on eigenvalue, which is a quantity representing the amount of variance in the data explained by the associated factor. A common rule of thumb is to keep the factors whose eigenvalues are greater than one.

Although, factor loadings provide clues regarding correlations between variables and factors, it is sometimes difficult to assign a variable to a factor when it is correlated with more than one factor. For example, the difficulty arises when many variables are correlated with the first factor, but some are positively correlated and others are negatively correlated, or when some of the variables load significantly to more than two factors at the same time.[21] In this situation, selected common factors are rotated by either orthogonal transformation or oblique transformation to enhance their interpretation. Although the oblique transformation results in correlated common factors, the two methods provide equal power of transformed factors. Hence, either transformation method may be used.

There are many factor extraction methods available in the Statistical Analysis System (SAS). They are the principal component analysis, principal factor analysis, iterated principal factor analysis, unweighted least-squares factor analysis, maximum-likelihood factor analysis, alpha factor analysis, image component analysis, and Harris component analysis.[16] Only the principal factor analysis and maximum-likelihood (ML) factor analysis[20] were used to develop this peer grouping model. The principal factor analysis is the most commonly used method of common factor analyses because of its computational simplicity. It is similar to the principal component analysis except for the use of prior communality estimates. The ML factor analysis has some merits as it has desirable asymptotic properties and allows statistical tests regarding the number of factors to choose. In addition, the ML factor analysis does not require a multi-variate normal distribution.[16] However, ML is a resource-intensive procedure, and sometimes, it does not converge within the specified number of iterations.

5.3. Cluster Analysis

Like factor analysis, cluster analysis is a data reduction technique. The difference is that the factor analysis finds correlations among variables and the cluster analysis finds correlations among observations, in this case the hospitals. The cluster analysis requires a measure of similarity to group observations, where the similarities are usually determined on the basis of Euclidean distances or some association measures. The objective of cluster analysis is to group similar or inter-related observations in the same cluster.

Cluster analyses are hierarchical partitioning.[20] The hierarchical methods are based on an assumption that there exists a hierarchical structure in the clusters, that is, a cluster is always determined by merging the most similar two clusters. Hence, at the beginning, all observations are distinct clusters and they are merged in an iterative process based on their similarity until all similar clusters are merged into one. On the other hand, the partitioning methods do not assume the relationship between the parent clusters and the subsequent cluster. It starts with an assumption regarding the number of clusters and assigning each observation to the most appropriate cluster. Next, each cluster is evaluated to make adjustments, and the process ends when no further adjustments are needed. Hence, it is likely that the original clusters and the last clusters do not have any structural relationships in this approach.

All clustering algorithms available in SAS are based on hierarchical methods. SAS includes average linkage, the centroid method, complete linkage, density linkage, ML method, single linkage, two-stage density linkage, and Ward's minimum-variance method.[16] These algorithms differ in the way the distances are calculated, and hence they may create different clusters regardless of the structure of a given data set. It is, therefore, important to choose the most appropriate clustering algorithms for an application. In this peer grouping model, Ward's method[13] was used as the clustering algorithm. This algorithm minimizes the within-cluster sum of squares while maximizing the between-cluster sum of squares, where the sum of squares is measured by the distance between clusters and/or observations. It tends to join two clusters with a small number of observations or hospitals and creates large clusters each with a similar number of observations or hospitals.

In hierarchical cluster analysis, it is necessary to determine when the clustering process stops. SAS provides general statistics for this purpose. They are R^2, cubic clustering criterion, pseudo F test, and the pseudo t^2 statistic.[16] Each measure is capable of indicating the optimal number of clusters and, if all of these measures indicate a certain number, it would be the ideal choice.

5.4. Bootstrapping

Bootstrapping is a resampling approach that allows one to make inference on population parameters using repeated sampling.[22] It is a particularly useful method when the estimate is difficult to obtain using the conventional computational methods. It is a non-parametric approach since it does not require any distributional assumptions to use. Bootstrapping is based on a series of bootstrap samples, each is selected from a population using the sampling with replacement approach. Based on each bootstrap sample, an estimate is obtained, and the

bootstrapping process creates a distribution of parameter estimates. Since each bootstrap sample contains a slightly different sample with some observations sampled more than once, the average of the distribution could be used as an estimate. If the variance of the distribution is small, it would indicate the stability of the estimate. Bootstrapping should be used with caution, however, since it may create a more biased estimate than computational methods. If each bootstrap sample is insufficient, it may result in an unreliable estimate.

Although SAS provides the criteria for stopping the clustering process (See Section 5.3), they are generally dependent on the shape of data and are very sensitive to outliers. They also tend to conflict with each other in indicating the optimal number of clusters. Bootstrapping was used as an alternative approach for clustering in developing this peer grouping model. The rationale for this approach was that, if there was a true cluster structure in the hospital data, the number would be likely to emerge if the estimation processes were repeated many times. Since a bootstrap sample would contain similar hospitals with the original data except for a small noise factor; that is, the portion of repeated samples; it was considered as an ideal sampling approach for this application. The main idea was to create a distribution of a test statistic representing the validity of clusters through a repeated sampling procedure and determine the optimal number of clusters by studying the distribution. Hence, the bootstrapping process was expected to indicate a valid as well as a reliable estimate of the number of clusters.

The test statistic (R) was based on measures that were widely used in the computer science field. It is composed of two elements -- a ratio (W_k) of within-cluster to between-cluster dispersions and a measure of compactness (M_k) of clusters for a k-cluster scheme.[23] The formulas for these elements are as follows, assuming k-cluster partition:

$$W_k = \frac{1}{k}\sum_{i=1}^{k} Max[\frac{S_i+S_j}{T_{ij}}], \qquad M_k = \sum_{k=1}^{k} S_k^2,$$

where

$$S_i = \sqrt{\frac{1}{N_i}\sum_{j=1}^{N_i}(x_{ij}-g_i)^2}, \qquad T_{ij} = |g_i-g_j|,$$

N_i = number of observations for cluster I, x_{ij} is the observation j in cluster I, g_i is the centroid for cluster I.

The test statistic was based on the weighted average of the W_k and M_k, i.e.,

$$R = a\frac{\Delta W_k}{|\Delta W|} + (1-a)\frac{\Delta M_k}{|\Delta M|},$$

where $\Delta W_k = W_k - W_{k+1}$, $\Delta M_k = M_k - M_{k+1}$, $|W_k|$ and $|M_k|$ are the norms of W_k and M_k, respectively, and the a is the weight factor, which was set at 0.75 in this project.

First, a ML factor analysis was conducted using the original data including the 20 hospital variables. Since one factor was found to be sufficient for the data, the factor scores were not rotated. After the initial factor analysis, an iterative procedure began. The number of clusters, or the clustering schemes, was set to range from 3 to 50. Next, for each clustering scheme, say k (k=3 to 50), the procedure was repeated 100 times. Using the factor scores as a data set, a bootstrap sample was selected so that the sample size would be the same as the original sample size, that is, 5,326 hospitals. The bootstrap sample differed from the original data since not all hospitals were sampled, and in some instances, a single hospital was represented more than once. Using the bootstrap sample, the $W_{k(1)}$ and $M_{k(1)}$ statistics were computed for the first iteration and stored for future use. At the end of the 100th repetition for the clustering scheme k, the W_k, M_k, and R values were determined and stored. The process was repeated for k+1 clustering scheme with 100 repetitions, until the clustering scheme k reached 50. At the end of the process, the distribution of R values was examined to determine which clustering scheme resulted in the best R value. Hence, the clustering scheme determined through this process represented the optimal number of clusters for the data.

The bootstrapping procedure described above was implemented in a SAS program. The program is composed of several macros. Hence, it can be easily modified to run different procedures by altering some of the macro parameters. Detailed discussions about the program and usage are provided in the accompanying User's Guide.

5.5. Face Validation

The clustering (peer grouping) models developed by the above methods were subjected to independent face validation by several volunteer Peer Review Organizations (PROs). These PROs had developed their own peer grouping methodologies, some very basic and subjective (See Section 3.2).

The data and the analytical methods used in developing the cluster models were briefly described to the PROs and the importance of validation was emphasized (See instructions to PROs in Appendix D). The PROs were required to assign the task of validation to those personnel who were most familiar with the characteristics of acute care hospitals in their state.

Each of the volunteer PROs received lists of the hospitals in their state peer grouped according to the two cluster models (See Section 6.4), the above mentioned instructions, and a standard response form (See Appendix D). The PRO validators were asked to use their personal knowledge and assess the appropriateness of the inclusion of each hospital in one of the clusters (peer groups), note any disagreements with the clustering of hospitals, and report the disagreements with their comments using the standard response form.

6. Results

6.1. Preliminary Analysis.

As stated above (See Section 4.4), 5,326 acute care hospitals participated in Medicare and claimed for services in 1994. Of these, 882 hospitals could not be linked to the AHA data set or their values for all of the 20 variables used in analyses were missing. Consequently, data on 4,444 hospitals were used in developing peer grouping models. Table 2 shows the numbers of hospitals for which information on each variable was available. The mean values for each variable and their standard deviations are also given in Table 2.

Data on 78 variables describing hospital characteristics were known for the 4,444 hospitals. On examination of these data, it was evident that most variables could be redundant because of apparent inter-correlations. The number of variables were reduced by (a) examining the variables for their property of reflecting concepts of the hospital environment as found in the literature and the pre-existing peer grouping methodology developed by PROs; (b) including variables which measured the changing environment such as managed care, hospital systems, and out-patients; (c) examining the frequency distributions of the values for categorical variables; (d) running a correlation analysis on the continuous variables to remove duplicative variables while limiting the number of domains represented by the variables; and (e) running a factor analysis. Categorical variables with skewed frequency distributions, that is, frequency of more than 95 percent in any level/category, were excluded from the study. Variables representing gerontology services were reduced to 18 which were important services, such as intensive care and medical care for cardiovascular diseases and cancer, relevant to Medicare population. These 18 variables were not highly correlated, but they all indicated gerontology services. Hence they were converted into one variable -- the "number of gerontology services."

As a result, 78 variables were reduced to 20 variables (Table 2) which were used in the cluster analyses. Most of these 20 variables were inter-related (Table 3). These correlations between paired variables resulted after removing their relationship with other variables given in Table 3. Significant correlations existed between the total number of patient days and adjusted patient days, the average daily census and the number of set-up and staffed beds, and between the total number (FTEs) of registered nurses (RN) and the total number (FTEs) of hospital employees. As expected, the occupancy rate was also correlated with the average daily census and the number of set-up and staffed beds.

6.2. First One Level Model

The variables were entered into the principal factor analysis program in SAS. Two significant factors were found. The eigenvalue of the first factor was 9.27 and it accounted for 77 percent of the variation. The second factor had an eigenvalue of 1.19 and accounted for 10 percent of the variation. The two factors together accounted for 87 percent of the variation.

Once the factors had been estimated, it was necessary to interpret them. This involved assigning each resulting factor a name that reflected the importance of the factor in predicting

Table 2. Mean values of 20 variables used in cluster analyses

Variable	Number*	Mean	Standard Deviation
Total Medicare patient discharges (CASES)	5326	2065	2258
Total number of patient days (TOTLOS)	5326	14646	18277
Adjusted patient days (ADJPAT)	5131	21356	24245
Number of set-up and staffed beds (BDTOT)	5131	173.7	176.7
Average daily census (ADC)	5131	109.2	134.3
Occupancy rate (OCCRAT)	5131	0.54	0.2
Number of gerontology services (SERV2)	4478	6.31	4.41
Case Mix Index (CMI)	5238	1.24	0.3
% Medicare discharges transferred out (TRANS)	5326	0.96	0.04
DRG entropy (ENTROPY)	5326	4.14	0.71
% Medicare discharges for which at least one procedure was coded (ORHADSUR)	5326	0.48	0.26
Participation in HMO (HMO)	4478	0.08	0.28
Participation in PPO (PPO)	4478	0.18	0.38
MSA population size (MSASZ)	5131	2.38	2.4
Total RN FTEs (FTERN)	5131	168.3	225.6
Total hospital FTEs (FTEH)	5126	687.8	901.2
Membership in a health system (MHS)	4470	0.42	0.49
Total residents/average daily census (RESADC)	5131	0.05	0.23
Accreditation by JCAHO (JCAHO)	5131	0.76	0.43
For-profit ownership (PROFIT)	5131	0.13	0.34

* Number of hospitals for which the variable's values were known

Table 3. Partial correlations between variables controlling all other variables

Variable	CASES	TOTLOS	ADJPAT	BDTOT	ADC	OCCRAT	SERV2	CMI	TRANS	ENTROPY	ORHADSUR	HMO	PPO	MSASZ	FTERN	FTEH	MHS	RESADC	JCAHO	PROFIT
Total Medicare patient discharges (CASES)																				
Total number of patient days (TOTLOS)	35																			
Adjusted patient days (ADJPAT)	14	84*																		
Number of set-up and staffed beds (BDTOT)	28	-11	-1																	
Average daily census (ADC)	-33	24	-6	88*																
Occupancy rate (OCCRAT)	13	-7	3	-48	58															
Number of gerontology services (SERV2)	12	-12	10	18	-13	8														
Case Mix Index (CMI)	9	14	-18	-2	-2	-2	27													
% Medicare discharges transferred out (TRANS)	6	-4	2	6	-3	5	16	21												
DRG entropy (ENTROPY)	2	-7	9	6	-9	12	29	12	-2											
% Discharges w one procedure (ORHADSUR)	-8	2	2	1	-2	5	5	17	13	5										
Participation in HMO (HMO)	0	1	-2	-5	3	2	-1	4	-2	2	1									
Participation in PPO (PPO)	-1	0	1	5	-3	-1	6	-2	1	0	-1	42								
MSA Population Size (MSASZ)	-10	4	1	8	-6	6	5	5	17	-8	15	8	-9							
Total RN FTE's (FTERN)	12	9	-14	1	3	-3	5	10	-3	2	0	0	-3	1						
Total Hospital FTE's (FTEH)	0	-12	16	6	11	-7	4	-1	0	-2	1	4	0	1	75*					
Membership in a Health System (MHS)	7	-2	-2	6	-3	1	7	5	-1	-5	1	1	-3	9	1	-3				
Total Residents/Average Daily Census (RESADC)	-8	4	-3	-6	2	-4	5	1	0	-7	7	2	-3	7	3	16	-3			
Accredidation by JCAHO (JCAHO)	1	-9	10	3	-3	4	4	4	3	21	17	-1	3	9	-3	0	6	-5		
For-profit Ownership (PROFIT)	-5	9	-8	6	-5	-7	-3	11	11	-6	-3	-3	2	8	0	-6	24	-3	6	

Printed values are multiplied by 100 and rounded to the nearest integer. Values greater than 3 are statistically significant, $p<0.01$

* Values greater than 60 indicate strong correlation

Table 4. Rotated factor loading (correlations between the factor and variables) in the first one-level model

Variable	Factor 1	Factor 2
Average daily census (ADC)	0.91*	0.23
Total Hospital FTE's (FTEH)	0.88*	0.33
Number of set-up and staffed beds (BDTOT)	0.87*	0.34
Total number of patient days (TOTLOS)	0.86*	0.37
Total RN FTE's (FTERN)	0.85*	0.36
Adjusted patient days (ADJPAT)	0.85*	0.37
Total Medicare patient discharges (CASES)	0.82*	0.44
Occupancy rate (OCCRAT)	0.48	0.13
Total Residents/Average Daily Census (RESADC)	0.27	0.08
Case Mix Index (CMI)	0.46	0.71*
Number of gerontology services (SERV2)	0.59	0.66*
% Medicare discharges transferred out of the hospital (TRANS)	0.35	0.66*
% Medicare discharges w procedures (ORHADSUR)	0.31	0.59
Accreditation by JCAHO (JCAHO)	0.18	0.56
DRG entropy (ENTROPY)	0.41	0.55
MSA Population Size (MSASZ)	0.31	0.54
Membership in a Health System (MHS)	0.06	0.40
For-profit Ownership (PROFIT)	-0.24	0.29
Participation in HMO (HMO)	0.08	0.10
Participation in PPO (PPO)	0.03	0.07

* Important variable for defining the Factor

each of the observed variables, that is, the coefficients in the pattern matrix corresponding to the factor. All the coefficients in a rotated pattern matrix were close to zero, or plus or minus 1. Therefore the standard rotation methods attempted to optimize a function of the pattern matrix that measured, in some sense, how close the elements were to zero, or plus or minus 1. The variables with the highest factor loadings were considered most important in defining the indicator. If the variables with the highest loadings together measured a concept (hospital size, hospital economics, diversity of services) then the factor was named.

Varimax rotation was used to help interpret the factors. The rotated factor pattern matrix is shown in Table 4. The variables with the highest factor loadings on Factor 1 -- average daily census, full-time hospital employees, number of set-up and staffed beds, etc. -- measure conceptual "hospital size" and "hospital efficiency." In Factor 2, the highest factor loadings -- case mix index, number of gerontology services, and the number of patients who were transferred out of the hospital -- seem to measure the conceptual "complexity of services." In summary, when all of the variables were in the factor analysis two significant factors emerged, that is, their eigenvalue was greater than 1.0. One measured the size and efficiency, and the other measured hospital complexity. Of the two, Factor 1 (size and efficiency) was more important, because it explained greater variation (See above).

The standardized factor scores were then used in Ward's minimum-variance cluster analysis. Using the pseudo statistics, it was determined that six clusters was the best solution. Table 5 shows the derived clusters with average values for the component variables. For example, Cluster 1 contained 736 hospitals which had averaged 325 Medicare patients, 1,653 total patient days, 2,947 adjusted patient days, and 1.47 gerontology services in 1994.

Table 5. Number of hospitals, average number of gerontology services, and patient days in each of the six clusters in the first one-level model

Cluster	Total Hospitals	Medicaid Cases	Patient Days	Adjusted Patient Days	Gerontology Services
1	736	324.6	1652.6	2946.7	1.5
2	904	696.9	3830.1	6644.2	3.2
3	1106	1853.4	12197.2	19166.5	6.1
4	895	2598.9	16856.8	23727.7	9.4
5	646	5065.8	39469.3	54877.3	10.7
6	157	8311.6	70638.8	91794.1	13.5

As previously described (See Section 2.2), the PROs for Indiana and Michigan had created peer grouping of hospitals in their respective states.[11,12] The staff of the two PROs face-validated this model by comparing the model's six clusters of hospitals in Indiana and Michigan with their previously developed state-specific peer groups. In addition, the staff at both PROs, who were not involved in the previous peer grouping work, provided their assessment. They concluded that the clusters did not seem to make sense. Hospitals seemed to be clustered on size and not on complexity. This cluster solution was considered inferior to the ones previously constructed in Indiana and Michigan. This model was rejected.

Bootstrapping of this model was conducted, and it was programmed to select 20 random samples testing 4 clusters through 30 clusters. The RSTAT statistic was the smallest (0.05) for the 26 cluster solution, suggesting that the most stable solution contained 26 clusters which were chosen by using pseudo-statistics, rather than the 6 clusters. This model was considered inferior to previous analyses because it seemed to simply differentiate hospitals on their size alone. These findings as well as poor face validity lead to its rejection.

6.3. Two-Level Methodology

The two-level methodology used for developing state-specific models for Indiana and Michigan was tried. It essentially stratifies the hospitals into clusters according to their size, and then the one-level model, described above, is applied to each stratum. By stratifying, the effect of hospital size is partially controlled at the outset. Thereby, the factor analysis produces factors which are more likely to measure complexity of hospital services and operational efficiency, that is, average daily census, staffing, etc.

6.3.1. First Two-Level Model

Two variables -- total length of stay and population size of metropolitan statistical area, taken as indicators of hospital and community size -- were entered into a cluster analysis which was forced to produce three major clusters. Each of these three clusters were then used to define a stratum. The first stratum included the small hospitals in the smallest communities. The second stratum included hospitals with less than 38,000 total patient days regardless of the community size. The third stratum included all of the larger hospitals. Factor analysis and cluster analysis was run on each of these three strata.

Factor analysis on the first stratum produced two significant factors. Factor 1 had an eigenvalue of 7.75 and accounted for 73 percent of the variation in the variables, and Factor 2 had an eigenvalue of 1.37 and accounted for 13 percent of the variation. Together, the factors explained 86.9 percent of the variation. The rotated factor pattern matrix is displayed in Table 6. Both factors measured various aspects of size and were difficult to name. The standardized factor scores were entered into a cluster analysis, and the pseudo statistics indicated that there were nine clusters in this stratum.

Factor analysis on the second stratum produced three significant factors. Factor 1 had an eigenvalue of 6.85 and accounted for 63.6 percent of the variation; Factor 2 had an eigenvalue

Table 6. Rotated factor loadings (correlations between the factor and variables), small hospitals in rural communities in the first two-level model

Variable	Factor 1	Factor 2
Total Medicare patient discharges (CASES)	0.87*	0.37
Adjusted patient days (ADJPAT)	0.84*	0.38
Total number of patient days (TOTLOS)	0.83*	0.42
Total RN FTE's (FTERN)	0.81*	0.45
Total Hospital FTE's (FTEH)	0.76*	0.53
Number of gerontology services (SERV2)	0.73*	0.27
DRG entropy (ENTROPY)	0.72*	0.19
Case Mix Index (CMI)	0.72*	0.17
Accreditation by JCAHO (JCAHO)	0.58	0.16
% Medicare discharges w procedures (ORHADSUR)	0.51	0.12
% Medicare discharges transferred out of the hospital (TRANS)	0.44	0.14
Membership in a Health System (MHS)	0.15	0.03
For-profit Ownership (PROFIT)	0.13	-0.50
Participation in PPO (PPO)	0.09	0
Total Residents/Average Daily Census (RESADC)	0.02	0
Participation in HMO (HMO)	0.02	-0.01
Average daily census (ADC)	0.13	0.97*
Number of set-up and staffed beds (BDTOT)	0.32	0.90*
Occupancy rate (OCCRAT)	0.04	0.52

* Important variable for defining the Factor
Note: MSA population size was not included in this model because of singular matrix problems since this stratum was the non-metropolitan cluster.

of 1.49 and accounted for 13.9 percent of the variation; and Factor 3 had an eigenvalue of 1.20 and it accounted for 11.2 percent of the variation. Together, the three factors explained 88.7 percent of the variation.

Table 7 shows the rotated factor pattern matrix. Factor 1 measured hospital operational efficiency, that is, the average daily census, total beds, and the number of full-time employees. Factor 2 measured hospital complexity, that is, case-mix index, the number of gerontology services, and percentage of patients transferred out of the hospital. Factor 3 measured the size in terms of the adjusted patient days, total length of stay, and the number of Medicare patients. The standardized factor scores were included in a cluster analysis, and the pseudo statistics indicated that there were eight clusters.

The factor analysis on the third stratum produced three significant factors. Factor 1 had an eigenvalue of 6.9 and accounted for 61.1 percent of the variation; Factor 2 had an eigenvalue of 1.63 and accounted for 14.4 percent of the variation; Factor 3 had an eigenvalue of 1.25 and it accounted for 11.1 percent of the variation. Together the three factors explained 86.7 percent of the variation.

Table 8 displays the rotated factor pattern matrix. The first factor measured size and operational efficiency -- total length of stay, adjusted patient days, number of Medicare patients, average daily census and total beds. The second factor measured complexity and staffing -- the number of services, case-mix index, full-time RNs, full-time staff, and the percentage of patients transferred out. The third factor measured occupancy rate which was also an indicator of operational efficiency. The standard scores were entered into a cluster analysis, and the pseudo statistics indicated six clusters.

Technical and program staff at the PROs for Indiana and Michigan tested the validity of these results, and again they were not acceptable. The clustering schemes of prior analyses did a better job of clustering hospitals. Bootstrapping of the first 2-level model indicated that there were 26 stable clusters. This model was rejected following face validation.

6.3.2. Second Two-Level Model

Two variables -- the total number of patient days and set-up and staffed beds -- were entered in Ward's cluster analysis which was required to produce three clusters which were used as strata. Each stratum was comprised of hospitals with similar characteristics. The first stratum included small hospitals defined by both total patient days and set-up and staffed beds, that is, hospitals with less than 21,882 patient days and less than approximately 200 beds. The second stratum included all medium size hospitals (200 to 550 beds and less than 54,705 patient days). The third stratum included all hospitals which had large numbers of beds and total number of patient days, that is, greater than 54,705 patient days or greater than 550 beds. A factor analysis and a cluster analysis were then run on each of these strata.

Factor analysis of small hospitals produced three factors. Factor 1 had an eigenvalue of 6.90 and accounted for 64.8 percent of the variation in the variables; Factor 2 had an eigenvalue of

Table 7. Rotated factor loadings (correlations between the factor and variables), medium size hospitals in the first two-level model

Variable	Factor 1	Factor 2	Factor 3
Average daily census (ADC)	0.94*	0.09	0.04
Number of set-up and staffed beds (BDTOT)	0.88*	0.30	0.07
Total Hospital FTE's (FTEH)	0.82*	0.37	0.15
Total RN FTE's (FTERN)	0.77*	0.43	0.14
Occupancy rate (OCCRAT)	0.54	-0.14	0.29
Total Residents/Average Daily Census (RESADC)	0.15	0.04	-0.14
Case Mix Index (CMI)	0.18	0.65*	0.23
Number of gerontology services (SERV2)	0.41	0.62*	0.39
% Medicare discharges transferred out of hosp. (TRANS)	0.24	0.61*	0.08
% Medicare discharges w procedures (ORHADSUR)	0.17	0.43	0.05
Membership in a Health System (MHS)	-0.02	0.41	-0.11
For-profit Ownership (PROFIT)	-0.26	0.37	-0.26
Accreditation by JCAHO (JCAHO)	0.07	0.27	0.17
Adjusted patient days (ADJPAT)	0.41	0.33	0.80*
Total number of patient days (TOTLOS)	0.44	0.39	0.77*
Total Medicare patient discharges (CASES)	0.38	0.44	0.75*
DRG entropy (ENTROPY)	0.20	0.32	0.50
Participation in PPO (PPO)	-0.01	0.00	0.22
Participation in HMO (HMO)	0.01	0.02	0.15
MSA Population Size (MSASZ)	0.03	0.13	-0.25

* Important variable for defining the Factor

Table 8. Rotated factor loadings (correlations between the factor and variables), large hospitals in the first two-level model

Variable	Factor 1	Factor 2	Factor 3
Total number of patient days (TOTLOS)	0.96*	0.09	0.06
Adjusted patient days (ADJPAT)	0.96*	0.03	0.05
Total Medicare patient discharges (CASES)	0.87*	0.19	-0.21
Average daily census (ADC)	0.72*	0.50	0.33
Number of set-up and staffed beds (BDTOT)	0.71*	0.55	0.05
Number of gerontology services (SERV2)	0.32	0.71*	-0.16
Case Mix Index (CMI)	0.22	0.69*	-0.18
Total RN FTE's (FTERN)	0.59	0.67*	0.18
Total Hospital FTE's (FTEH)	0.60*	0.65*	0.23
% Medicare discharges transferred out of hosp. (TRANS)	0.16	0.64*	-0.31
DRG entropy (ENTROPY)	0.23	0.60	0.10
Total Residents/Average Daily Census (RESADC)	0.08	0.57	0.35
% Medicare discharges w procedures (ORHADSUR)	0.12	0.39	0.24
Participation in HMO (HMO)	-0.05	0.19	0.15
Accreditation by JCAHO (JCAHO)	0.05	0.09	0.01
Participation in PPO (PPO)	-0.01	0.04	0.01
Occupancy rate (OCCRAT)	0.22	-0.01	0.73*
MSA Population Size (MSASZ)	-0.02	0.08	0.24
For-profit Ownership (PROFIT)	-0.12	-0.06	-0.21
Membership in a Health System (MHS)	0.11	0.10	-0.27

* Important variable for defining the Factor

1.59 and accounted for 14.9 percent of the variation; and Factor 3 had an eigenvalue of 1.13 and accounted for 10.6 percent of the variation. Together the factors explained 90.3 percent of the variation.

Table 9 shows the rotated factor pattern matrix for small hospitals. The first factor measured various aspects of size -- adjusted patient days, the number of Medicare cases, total number of patient days, entropy, and to a lesser degree staffing. The weightings of the second factor are much smaller, but it measured aspects of staffing and complexity of services, that is, the number of gerontology services and case-mix index. The third factor measured hospital efficiency -- average daily census, occupancy rate and bed total. The standardized factor scores were entered in a cluster analysis, and the pseudo statistics indicated that there were nine clusters in this stratum.

The factor analysis for medium size hospitals produced three significant factors. Factor 1 had an eigenvalue of 6.2 and accounted for 57.1 percent of the variation; Factor 2 had an eigenvalue of 1.7 and accounted for 16 percent of the variation, and Factor 3 had an eigenvalue of 1.5 and accounted for 13.5 percent of the variation. Together the three factors explained 86.6 percent of the variation.

Table 10 shows the rotated factor pattern matrix. Factor 1 measured complexity -- staffing and services. Factor 2 measured size -- adjusted patient days, total length of stay, and the number of Medicare cases. Factor 3 measured hospital operational efficiency, that is, the average daily census and occupancy rate. The standardized factor scores were entered in a cluster analysis, and the pseudo statistics indicated that there were 10 clusters.

The factor analysis for large hospitals produced four significant factors. Factor 1 had an eigenvalue of 5.4 and accounted for 46.2 percent of the variation; Factor 2 had an eigenvalue of 2.5 and accounted for 21.3 percent of the variation; and Factor 3 had an eigenvalue of 1.7 and accounted for 14.5 percent of the variation. Factor 4 had an eigenvalue of 1.1 and accounted for 9.8 percent of the variance. Together, the first three factors and all four factors explained 82.1 percent 91.9 percent of the variation respectively. Since the fourth factor was only marginally significant, it was dropped from the analysis.

Table 11 displays the rotated factor pattern matrix. The first factor measured hospital complexity -- the number of gerontological services, entropy, case mix index, and transfers out. The second factor measured operational efficiency, that is, staffing and occupancy rate. The third factor measured size -- adjusted patient days, total length of stay, and number of cases. The standard scores were entered in a cluster analysis, and the pseudo statistics indicated 10 clusters.

The program and technical staff at the PROs for Indiana and Michigan tested the face validity of the results. Of all the models tested heretofore, this was the model which had the best face validity. However, the clusters did not discriminate well on size. Hospitals of very diverse sizes in terms of the number of cases, the number of set-up and staffed beds, average daily census, etc. were in the same cluster. It was nearly impossible to define the cluster.

Table 9. Rotated factor loadings (correlations between the factor and variables), small hospitals in the second two-level model

Variable	Factor 1	Factor 2	Factor 3
Adjusted patient days (ADJPAT)	0.91*	0.16	0.15
Total Medicare patient discharges (CASES)	0.91*	0.22	0.19
Total number of patient days (TOTLOS)	0.90*	0.20	0.21
Total Hospital FTE's (FTEH)	0.67*	0.49	0.25
Total RN FTE's (FTERN)	0.63*	0.55	0.25
DRG entropy (ENTROPY)	0.61*	0.09	0.12
Number of gerontology services (SERV2)	0.55	0.42	0.16
Accreditation by JCAHO (JCAHO)	0.42	0.40	0.08
Participation in PPO (PPO)	0.11	-0.03	-0.01
Participation in HMO (HMO)	0.06	0.00	0.01
MSA Population Size (MSASZ)	0.12	0.55	0.06
Case Mix Index (CMI)	0.39	0.55	0.06
% Medicare discharges transferred out of hosp. (TRANS)	0.21	0.48	0.09
% Medicare discharges w procedures (ORHADSUR)	0.25	0.47	0.08
For-profit Ownership (PROFIT)	-0.05	0.44	-0.01
Membership in a Health System (MHS)	0.01	0.38	0.03
Total Residents/Average Daily Census (RESADC)	-0.06	0.11	-0.04
Average daily census (ADC)	0.21	0.14	0.95*
Occupancy rate (OCCRAT)	0.11	-0.16	0.80*
Number of set-up and staffed beds (BDTOT)	0.33	0.43	0.71*

* Important variable for defining the Factor

Table 10. Rotated factor loadings (correlations between the factor and variables), medium size hospitals in the second two-level model

Variable	Factor 1	Factor 2	Factor 3
Total RN FTE's (FTERN)	0.82*	0.25	0.21
Total Hospital FTE's (FTEH)	0.81*	0.23	0.29
Number of set-up and staffed beds (BDTOT)	0.79*	0.15	0.29
Average daily census (ADC)	0.70*	0.10	0.63*
Number of gerontology services (SERV2)	0.66*	0.37	-0.15
Case Mix Index (CMI)	0.65*	0.28	-0.27
% Medicare discharges transferred out of hosp. (TRANS)	0.61*	0.28	-0.29
Total Residents/Average Daily Census (RESADC)	0.44	-0.09	0.14
% Medicare discharges w procedures (ORHADSUR)	0.42	0.12	-0.04
MSA Population Size (MSASZ)	0.36	0.03	-0.07
Participation in HMO (HMO)	0.14	-0.01	0.01
Adjusted patient days (ADJPAT)	0.19	0.89*	0.33
Total number of patient days (TOTLOS)	0.27	0.88*	0.32
Total Medicare patient discharges (CASES)	0.32	0.86*	0.16
DRG entropy (ENTROPY)	0.38	0.41	-0.10
Accreditation by JCAHO (JCAHO)	0.02	0.12	-0.02
Participation in PPO (PPO)	-0.01	0.02	-0.01
Occupancy rate (OCCRAT)	0.09	0.06	0.78*
Membership in a Health System (MHS)	0.21	0.04	-0.28
For-profit Ownership (PROFIT)	-0.03	-0.05	-0.42

* Important variable for defining the Factor

Table 11. Rotated factor loadings (correlations between the factor and variables), large hospitals in the second two-level model

Variable	Factor 1	Factor 2	Factor 3
Number of gerontology services (SERV2)	0.79*	0.11	0.23
DRG entropy (ENTROPY)	0.69*	-0.11	0.29
Case Mix Index (CMI)	0.65*	0.04	0.24
% Medicare discharges transferred out of hosp. (TRANS)	0.60*	0.00	0.20
Total Residents/Average Daily Census (RESADC)	0.51	0.39	-0.12
% Medicare discharges w procedures (ORHADSUR)	0.48	0.07	0.12
Accreditation by JCAHO (JCAHO)	0.46	-0.19	0.16
Participation in HMO (HMO)	0.17	0.04	-0.02
Average daily census (ADC)	-0.03	0.96*	0.04
Number of set-up and staffed beds (BDTOT)	0.11	0.88*	0.07
Total Hospital FTE's (FTEH)	0.56	0.71*	0.18
Total RN FTE's (FTERN)	0.60	0.63*	0.18
Occupancy rate (OCCRAT)	-0.24	0.42	0.02
MSA Population Size (MSASZ)	0.07	0.27	0.04
Participation in PPO (PPO)	0.05	-0.08	0.05
For-profit Ownership (PROFIT)	-0.01	-0.11	-0.04
Adjusted patient days (ADJPAT)	0.13	0.17	0.95*
Total number of patient days (TOTLOS)	0.16	0.19	0.95*
Total Medicare patient discharges (CASES)	0.28	0.06	0.85*
Membership in a Health System (MHS)	0.04	0.00	0.13

* Important variable for defining the Factor

Bootstrapping results indicated that there were 30 stable clusters. When the size was controlled, other factors such as complexity and operational efficiency played important roles. But the balance between size and complexity/efficiency/staffing again was not achieved. The clusters showed considerable spread. Consequently, it was decided to try other factor analysis and clustering methods.

6.4. Second One-Level Methodology

The principal factor analysis is the most common factor analysis technique. An alternative method, the maximum likelihood (ML) method, is preferred by some, has desirable asymptotic properties, gives better estimates for large samples, and it does not require a multi-variate normal distribution. It is also suggested that principal factor analysis be used to get a rough idea of the relationship among variables, and then the maximum likelihood could be applied to refine the analysis. ML factor analysis does have its draw backs. It is very resource intensive because it uses an iterative process to calculate the factors, it cannot be used with a singular correlation matrix, and it is especially prone to Heywood cases.

The same variables, used in the above models, were entered in the ML factor analysis program. The model converged in 16 iterations. The model was statistically significant; Chi-square = 94432, df = 190, and $p<0.0001$. One significant factor was found with an eigenvalue of 150.2 and it accounted for 99 percent of the variation in the variables. The squared canonical correlation was 0.99. Because only one factor was needed, rotation was not necessary. Table 12 displays the resulting factor.

The standardized scores were entered into Ward's cluster analysis. The bootstrapping program ran 100 random samples for each cluster solution containing from three to 35 clusters. The results demonstrated that 23 clusters was the most stable solution. The program was re-run to generate an additional 100 random samples. Once again, the solution was 23 stable clusters. Extensive face validation was conducted on this final model with 23 clusters (See Section 6.5).

The average values of six major variables for each of the 23 clusters are given in Table 13, and their standard deviations and the ranges of values are given in Appendix E. For example, Cluster "A" contained 1,002 hospitals which had on average 263 patients in 1994 who stayed in the hospital a total of 1,223 days (total patient days). The 2,281 adjusted patient days included outpatient activity. The average total hospital beds were 51, with an average daily census of 27.5, and an occupancy rate of 43 percent. Hospitals in this cluster on average provided two gerontological services. On the other hand, cluster "W"contained three hospitals with average number of Medicare cases of 14, 349, the total length of stay 151,237, etc.

Figures 1 through 6 show the relationships between the Factor (the standardized scale) and key variables. All of these correlations are statistically very significant ($p<0.0001$). Each alpha character in the scatter plots denotes one or more hospitals in a cluster. One cluster may superimpose another cluster, and an alpha character may totally conceal another alpha character. Figure 1 displays the relationship between the Factor and the total number of set-up and staffed beds. The clusters for the most part are quite distinct; that is, "A's" are together,

Table 12. Rotated factor loadings (correlations between the factor and variables) in the second one-level model

Variable	Factor 1
Total number of patient days (TOTLOS)	0.99*
Adjusted patient days (ADJPAT)	0.99*
Total Medicare patient discharges (CASES)	0.97*
Number of set-up and staffed beds (BDTOT)	0.84*
Total Hospital FTE's (FTEH)	0.84*
Total RN FTE's (FTERN)	0.84*
Average daily census (ADC)	0.82*
Number of gerontology services (SERV2)	0.76*
Case Mix Index (CMI)	0.67*
DRG entropy (ENTROPY)	0.58
% Medicare discharges transferred out of hosp. (TRANS)	0.55
% Medicare discharges w procedures (ORHADSUR)	0.49
MSA Population Size (MSASZ)	0.47
Occupancy rate (OCCRAT)	0.45
Accreditation by JCAHO (JCAHO)	0.38
Total Residents/Average Daily Census (RESADC)	0.21
Membership in a Health System (MHS)	0.21
Participation in HMO (HMO)	0.09
Participation in PPO (PPO)	0.06
For-profit Ownership (PROFIT)	-0.08

* Important variable for defining the Factor

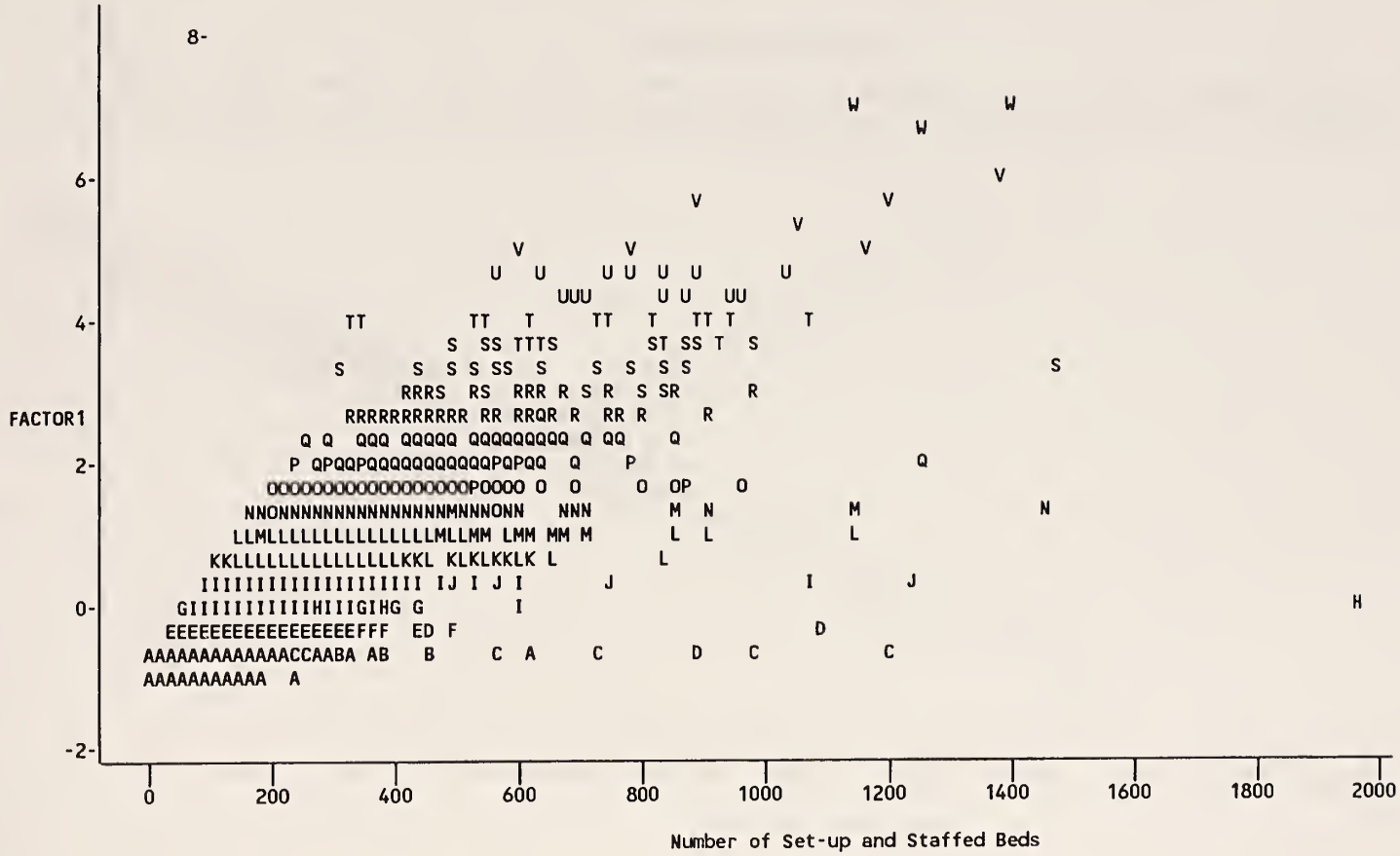
Table13. Number of hospitals and patients with average values of six variables according to each cluster

Cluster	Total Hospitals	Total Medicare Cases	Total Patient Days	Adjusted Patient Days	Set-up & Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
Missing	882	1,330	9,251.1	13,256.3	140.2	84.0	51%	3
A	1002	263	1,223.2	2,280.6	50.8	27.5	43%	2
B	450	835	4,505.8	7,705.3	89.9	49.6	49%	4
C	368	596	3,041.2	5,419.9	65.6	34.9	47%	3
D	250	1,114	6,316.6	10,421.6	105.9	56.5	50%	5
E	306	1,403	8,438.4	13,453.5	120.7	63.6	52%	6
F	187	1,707	10,468.3	16,544.4	144.7	79.2	54%	7
G	264	2,031	12,970.6	19,935.3	160.6	89.8	56%	7
H	174	2,358	15,836.2	23,973.4	197.3	119.2	59%	8
I	218	2,796	18,960.4	27,730.7	222.7	138.9	62%	9
J	195	3,215	22,333.0	32,826.2	255.3	158.4	62%	10
K	240	3,715	26,487.8	38,140.1	277.4	176.7	64%	10
L	126	4,229	30,566.9	43,451.8	326.4	211.7	65%	11
M	155	4,735	35,259.6	49,391.6	351.2	232.5	66%	11
N	99	5,377	39,724.6	55,002.4	373.7	256.6	69%	12
O	132	5,780	45,057.1	61,714.8	391.4	268.3	70%	12
P	70	6,522	52,076.9	68,724.4	449.0	323.4	73%	12
Q	81	7,298	58,454.1	78,099.3	532.4	381.4	72%	13
R	51	8,021	68,939.9	90,891.4	574.6	423.0	74%	13
S	33	9,683	80,826.1	104,658.3	692.7	528.4	77%	14
T	17	11,259	92,807.9	119,126.4	723.8	546.1	77%	15
U	14	12,024	103,524.5	134,781.4	798.6	610.4	76%	15
V	9	13,468	120,623.3	157,204.4	1,016.6	735.2	70%	15
W	3	14,349	151,237.0	199,863.8	1,269.7	1,022.3	81%	15

Table 14. Number of hospitals and patients with average values of six variables according to each peer group

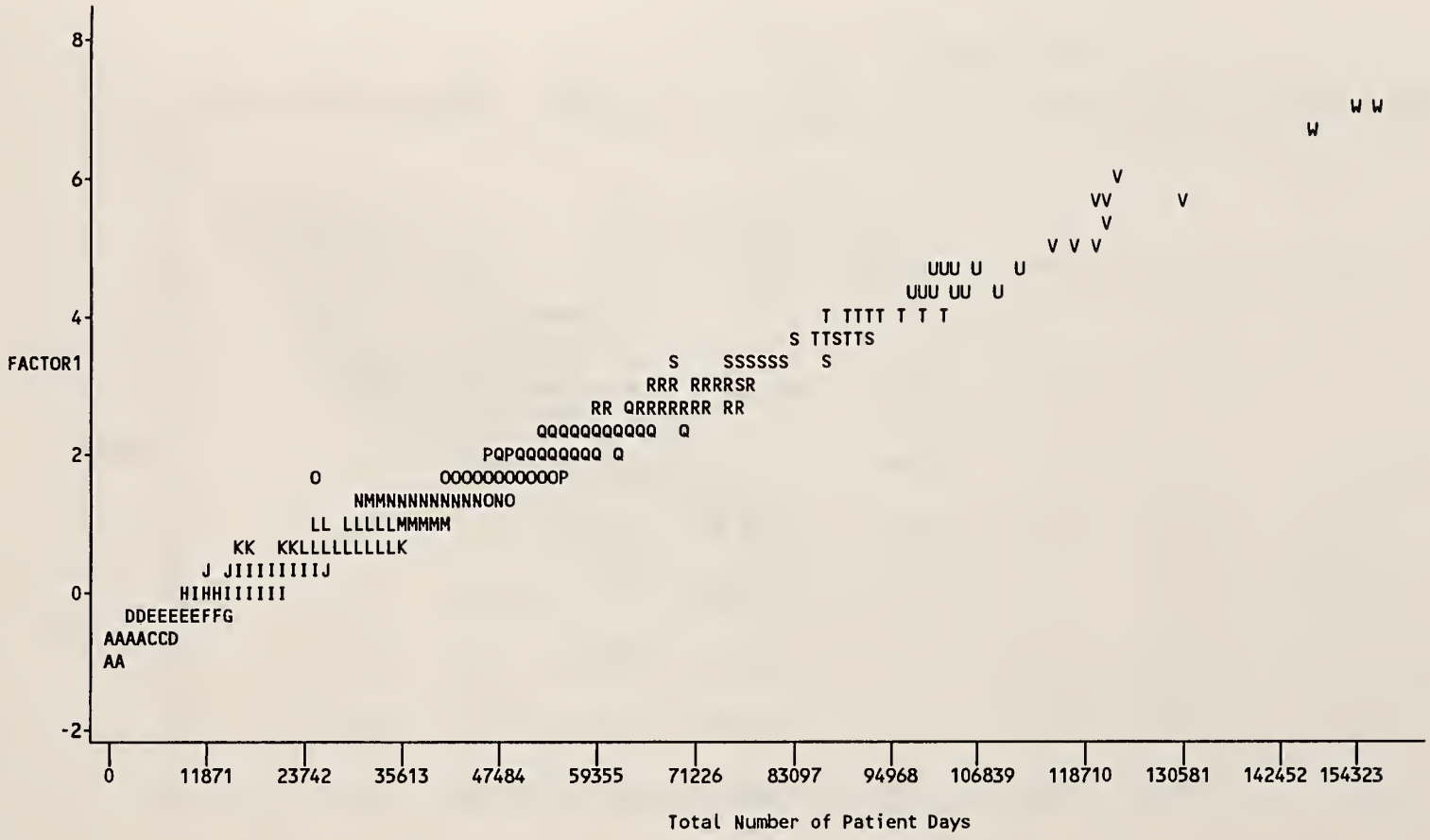
Peer Group	Total Hospitals	Total Medicare Cases	Total Patient Days	Adjusted Patient Day	Set-up Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
A	2070	549	2,875.2	5,001.2	68.6	37.1	47%	3
B	757	1,697	10,520.4	16,477.5	140.5	76.6	54%	6
C	587	2,806	19,154.7	28,309.7	226.0	139.6	61%	9
D	366	3,892	27,892.1	39,968.7	294.3	188.8	65%	10
E	386	5,257	39,755.2	55,044.8	370.7	250.9	69%	12
F	235	7,559	61,971.0	81,812.0	539.2	393.8	74%	13
G	43	12,186	106,195.3	137,826.0	847.5	639.8	76%	15

Figure 1. Relationship between Factor 1 and the number of set-up and staffed beds.



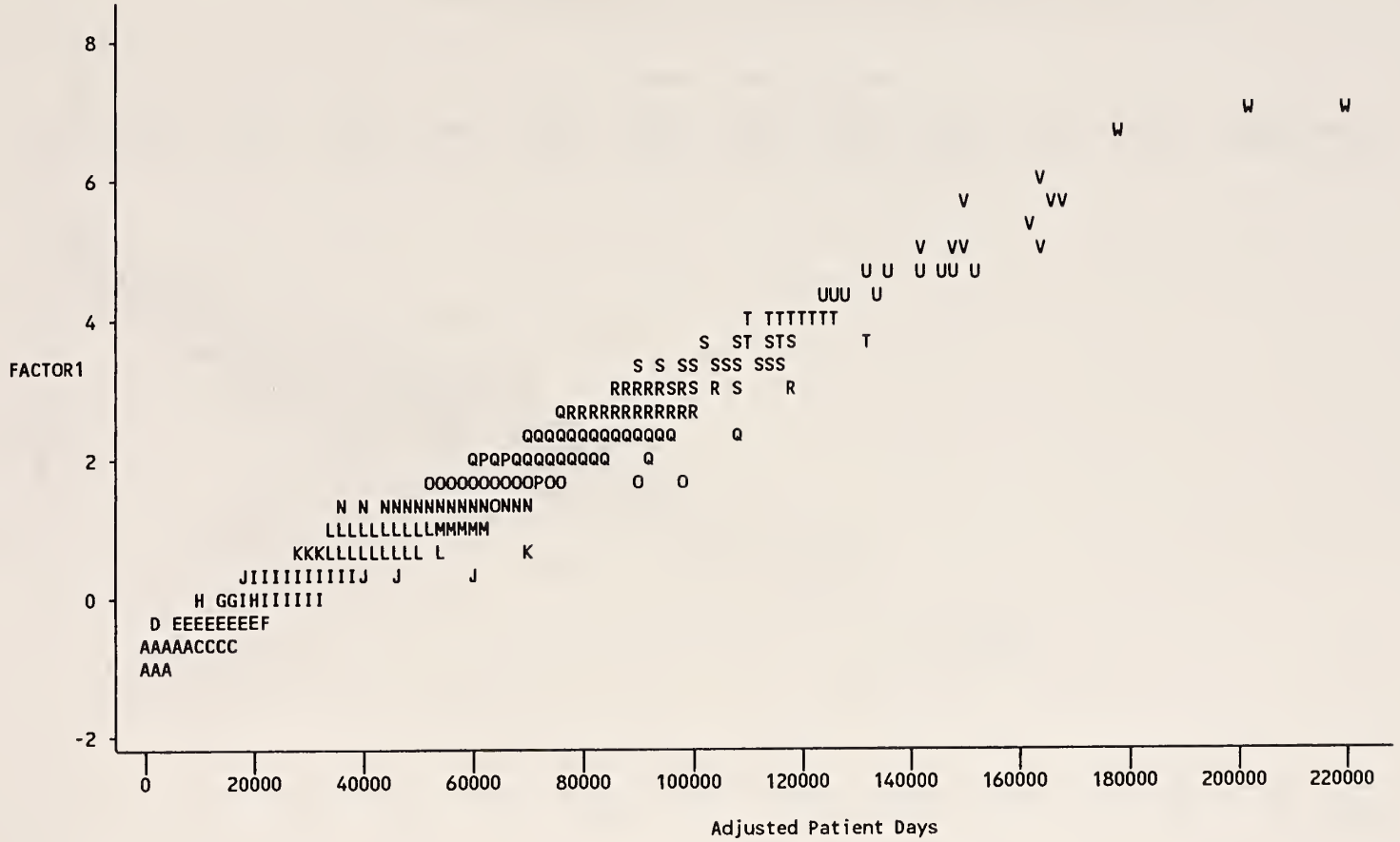
Alpha characters represent one or more hospitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.85$, $p<0.0001$

Figure 2. Relationship between Factor 1 and the total number of patient days.



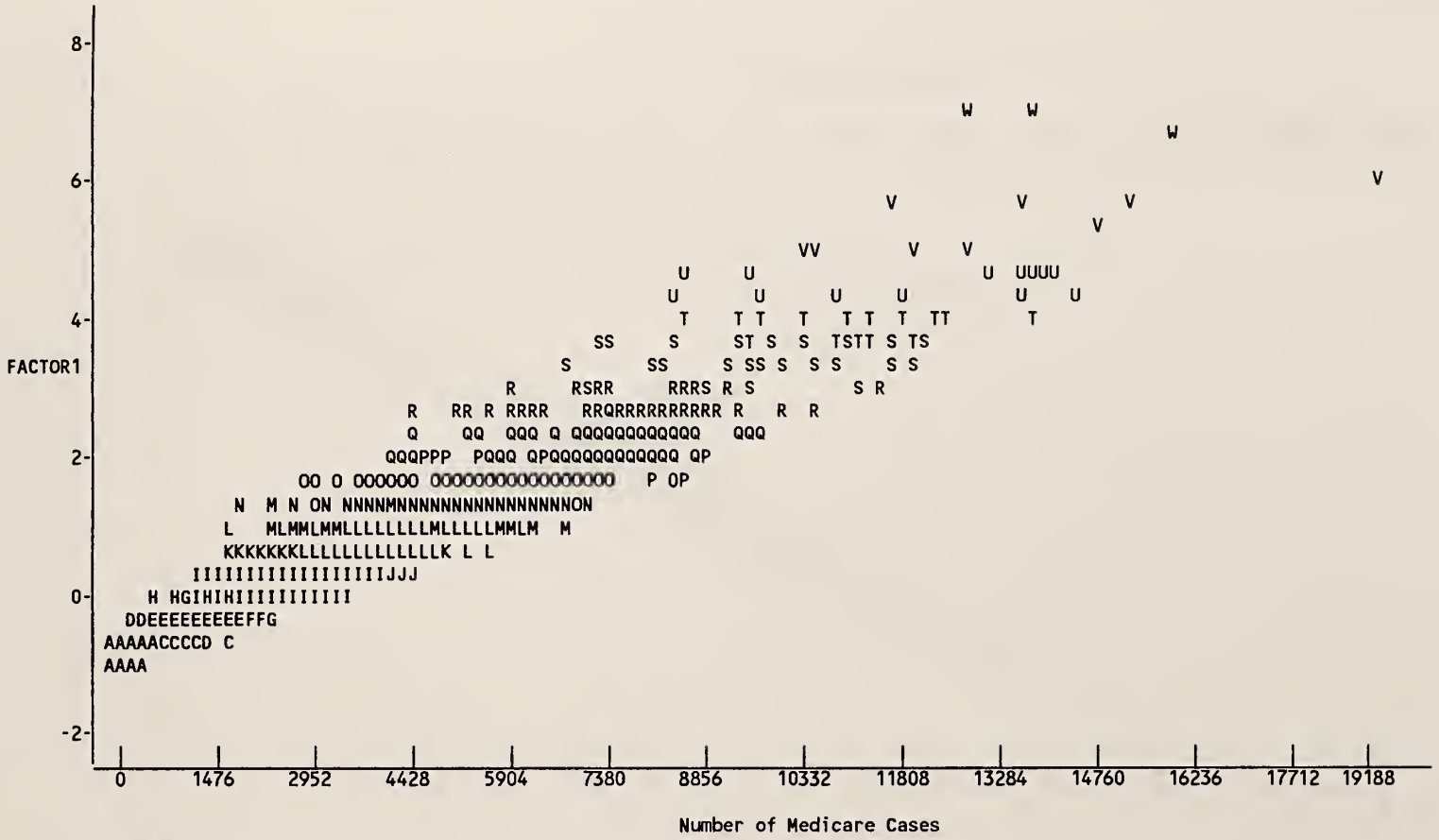
Alpha characters represent one or more hospitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.99$, $p<0.0001$

Figure 3: Relationship between Factor 1 and adjusted patient days.



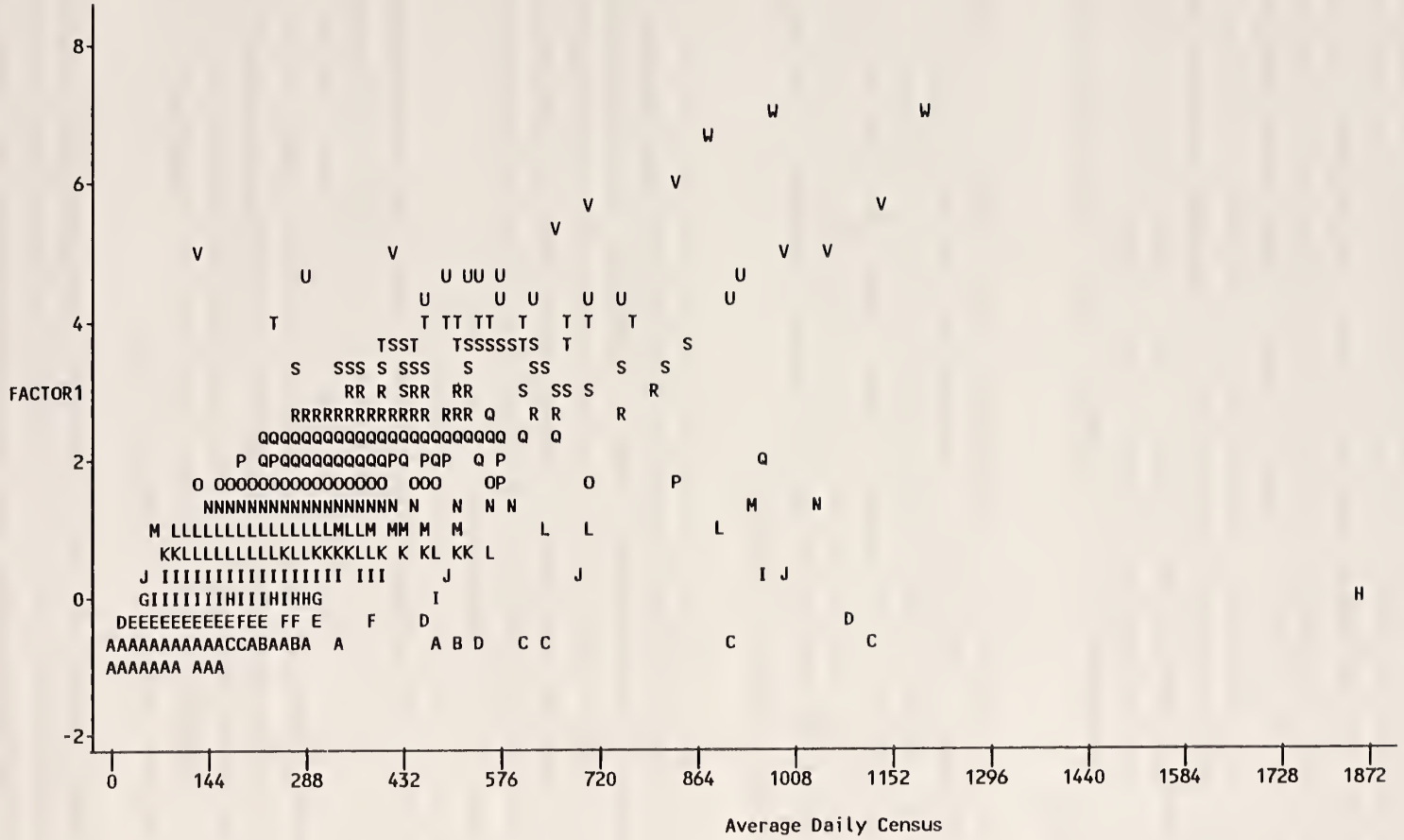
Alpha characters represent one or more hospitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.99$, $p<0.0001$

Figure 4. Relationship between Factor 1 and the number of Medicare cases.



Alpha characters represent one or more hospitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.97$, $p<0.0001$

Figure 5. Relationship between Factor 1 and average daily census.



Alpha characters represent one or more hopitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.83$, $p<0.0001$

"B's" are together, etc. There are, however, several out-lying (away from a diagonal band) scatter points. This suggests that the Factor is moderately related to the number of beds. The coefficient of correlation (r^2) was 0.85. Figure 2 shows a much stronger linear relationship between the Factor and the number of patient days ($r^2 = 0.99$). This suggests that the total number of patient days was a strong predictor of Factor 1. Figure 3 also shows that the adjusted patient days (total length of stay adjusted for outpatient activity) were strongly related to the Factor ($r^2 = 0.99$). The scatter shown in Figure 4 is slightly more spread, but it is still a very linear pattern of relatedness between the number of Medicare cases and the Factor ($r^2 = 0.97$). Figure 5 displays the relationship between the average daily census and the Factor. It is also more spread and there are more outliers, $r^2 = 0.83$). Figure 6 displays a relatively poor relationship between the Factor and the number of gerontological services ($r^2 = 0.76$). This variable had the most spread, the least predictability, and it contributed relatively little to the factor weighting.

Figures 7, 8, and 9 show inter-relationships between selected variables. The coefficients of partial correlations are given in Table 3. The relationship between the number of Medicare cases and the total number of beds is represented by the scatter plot in Figure 7. Although there were distinct clusters, there were also some outliers. Figure 8 shows the relationship between the number of Medicare cases and the total number of patient days. This appears to be a close relationship. This is also true of the relationship between the number of Medicare cases and the number of adjusted patient days (Figure 9).

In summary, the factor weightings, the descriptions of the clusters, and the scatter plots describe the final analysis. The clusters are based on a single hospital factor that is highly related to a number of hospital variables. It is most related to patient days and the number of adjusted patient days, somewhat related to the number of set-up and staffed beds and the average daily census, and weakly related to the number of gerontology services. Each of the final clusters remain quite distinct in the plots and in the analysis, which leads to the conclusion that this 23 cluster solution is strong.

The major practical weakness of the 23-cluster model was that there were too many clusters. A smaller number of clusters may be desirable for certain applications. Consequently, further clustering (grouping) of the 23 clusters was attempted using the Cluster Analysis in SAS. It was decided to collapse the 23 clusters into seven large clusters or peer groups. The resulting seven peer groups (large clusters) are given in Table 14. Peer Group A had 2,070 hospitals, and on average, 549 Medicare cases in 1994; whereas Peer Group G had only 43 hospitals and an average of 12,186 Medicare cases in 1994. Table 15 shows the relationship between the 7 peer groups (large clusters) and the 23 original clusters.

Figures 10 through 15 display the scatter plots of the seven large clusters or peer groups in terms of the relationships between Factor 1 and the key variables. The coefficients of these correlations and their statistical significance have been noted above. Figure 10 shows the relationship of the number of set-up and staffed beds with Factor 1. The relationship was linear with some variance (some outliers). Figures 11 and 12 show the relationship of Factor 1 with the total number of patient days and adjusted patient days respectively. In both graphs, the

Table 15. Composition of 7 large clusters (peer groups) in terms of the 23 clusters of hospitals

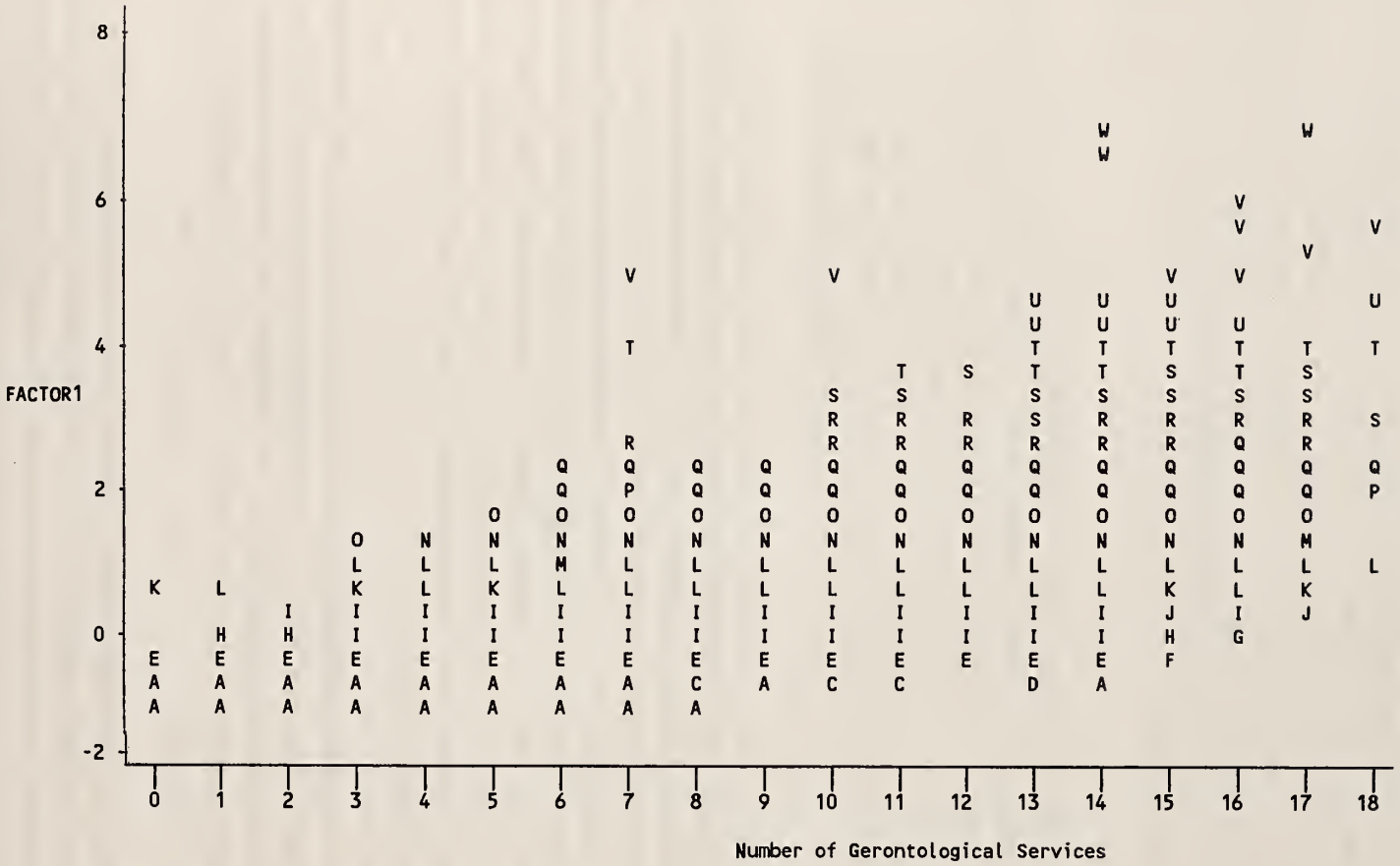
Large Clusters	23 Clusters
A	A, B, C, D
B	E, F, G
C	H, I, J
D	K, L
E	M, N, O
F	P, Q, R, S
G	T, U, V, W

relationship appears to be linear and very close, and the clusters do not overlap each other. The relationship between Factor 1 and the total number of Medicare cases is closer in large clusters A, B, and C (essentially small hospitals) than in other large clusters (Figure 13). The average daily census was linearly related to Factor 1, but there was some spread of hospitals within each large cluster, particularly the small hospitals (Figure 14). Figure 15 shows the relationship between Factor 1 and the total number of gerontological services. This relationship was not close -- considerable spread of hospitals within each cluster -- in both the 23-cluster and 7-cluster models (Figures 6 and 15). The number of gerontological services as a variable had only moderate predictability.

These results show that this single factor clustering method, which used the maximum likelihood factor analysis and Ward's cluster analysis, was the best model. The one strong factor (scale) in this model predicted 99 percent of the variation of the variables. Bootstrapping confirmed that 23 clusters was a stable solution. Validity tests demonstrated that only a few hospitals were misplaced.

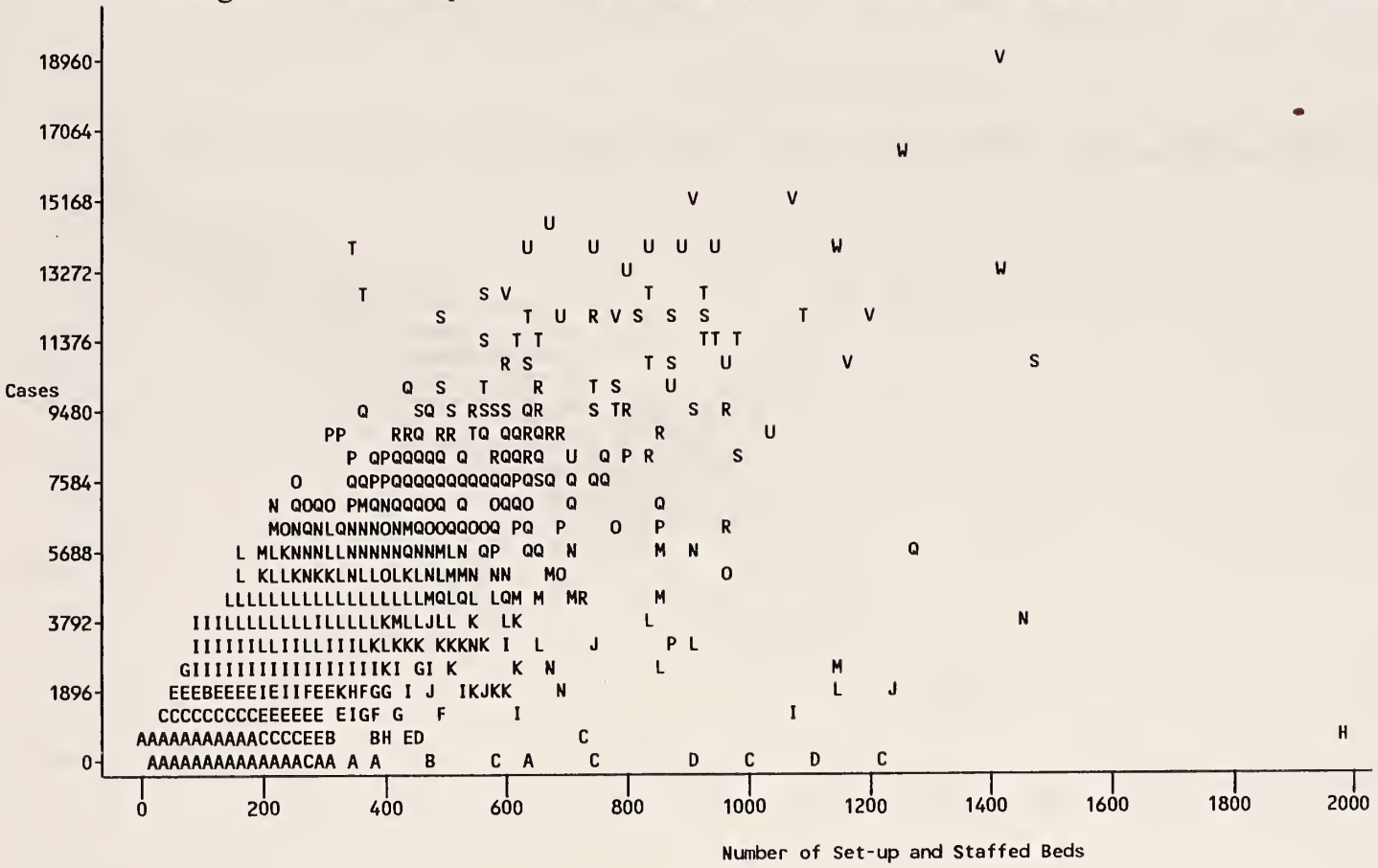
The 23-cluster model was collapsed to provide a 7-cluster peer grouping scheme. That is, the average factor weight of each of the 23 clusters was entered into a secondary cluster analysis and a seven large cluster scheme was requested. The results show that the pattern was very similar to that of the 23 cluster model. Thus the users will be provided with a choice of two clustering schemes -- 23 and seven (peer) groups of hospitals. Both models should provide similar grouping strategy.

Figure 6. Relationship between Factor 1 and the number of gerontology services.



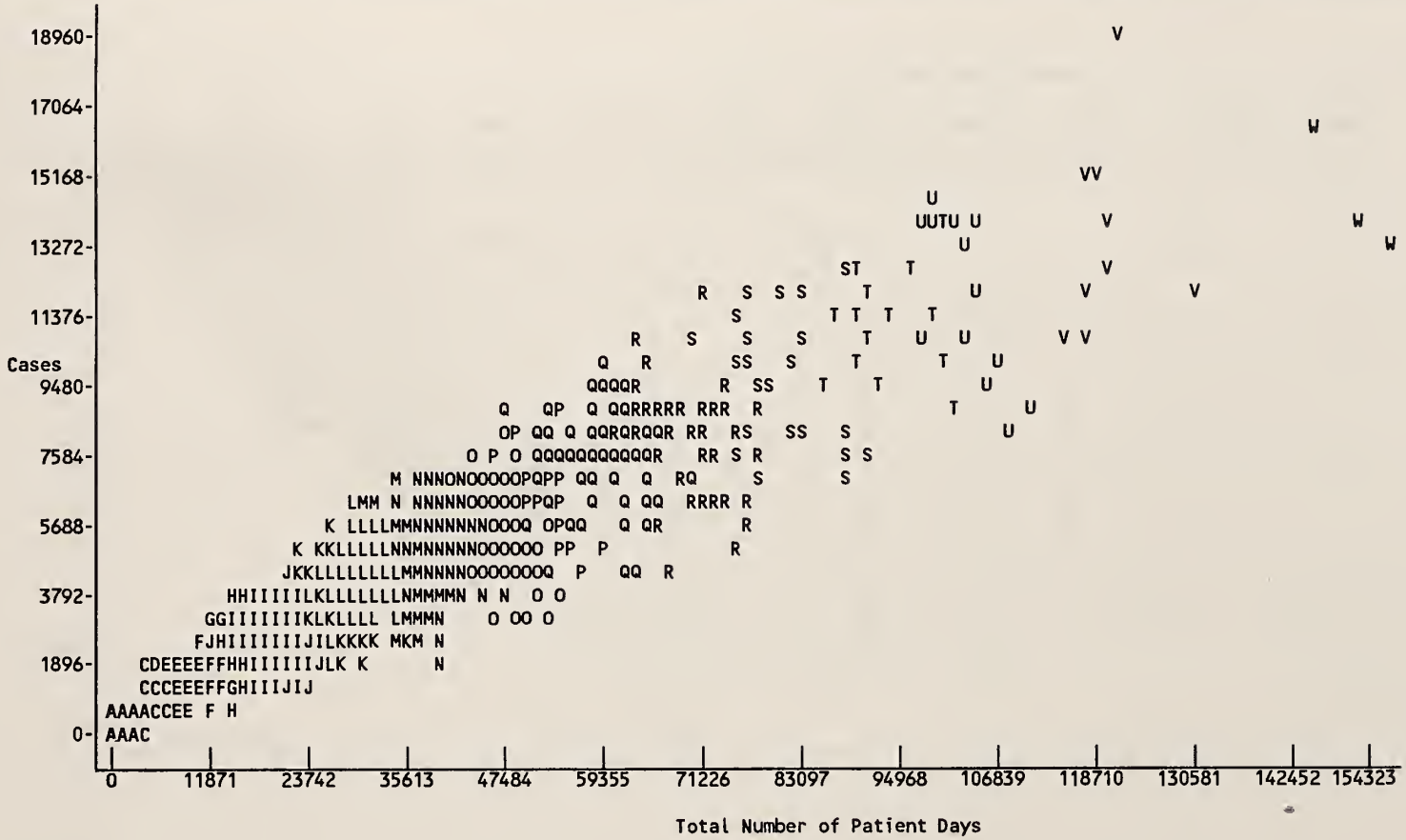
Alpha characters represent one or more hopitals in each of the 23 clusters. One cluster may superimpose another cluster.
$R^2 = 0.76$, $p<0.0001$

Figure 7. Relationship between the number of Medicare cases and the number of set-up and staffed beds.



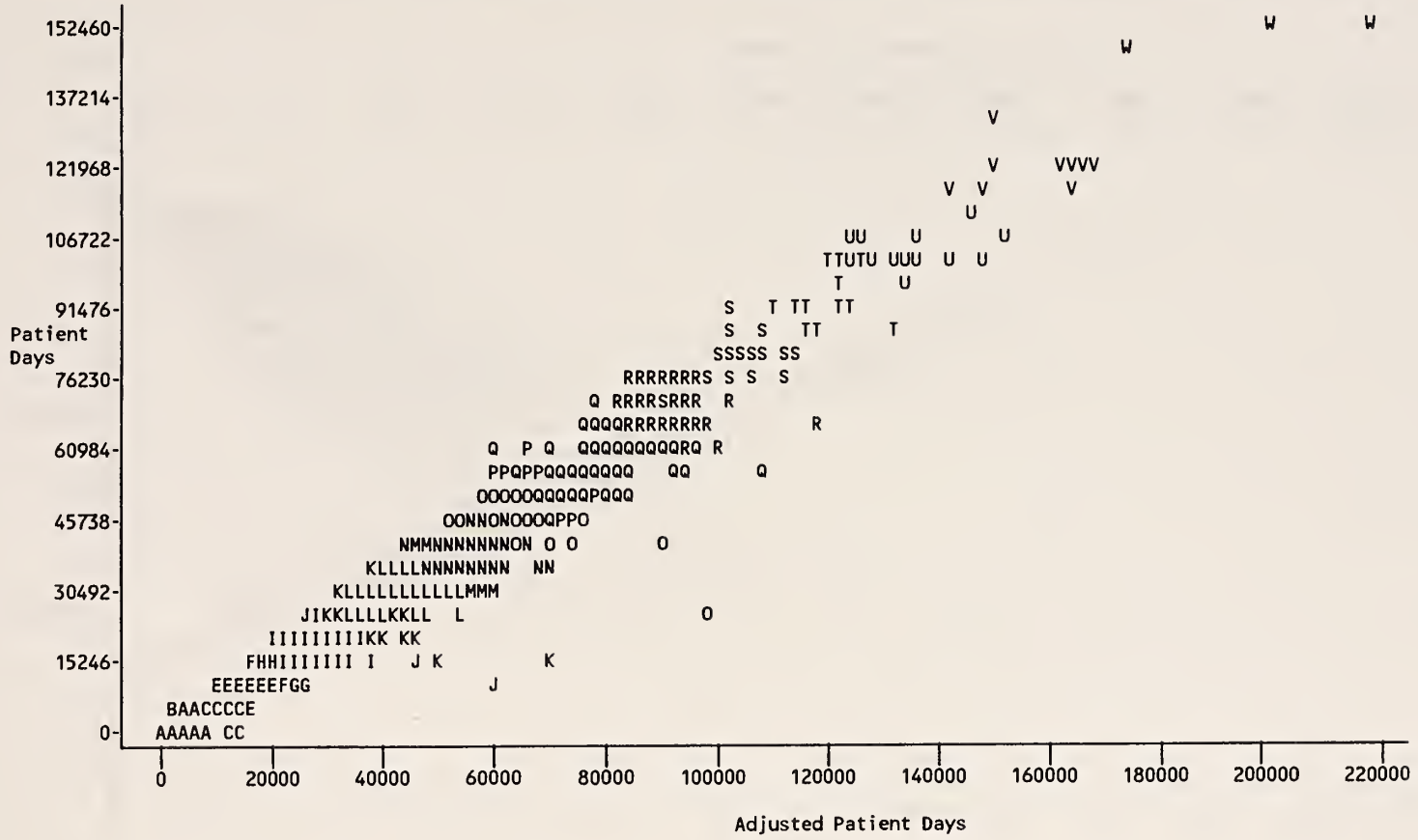
Alpha characters represent one or more hopitals in each of the 23 clusters. One cluster may superimpose another cluster.

Figure 8. Relationship between the number of Medicare cases and the total number of patient days.



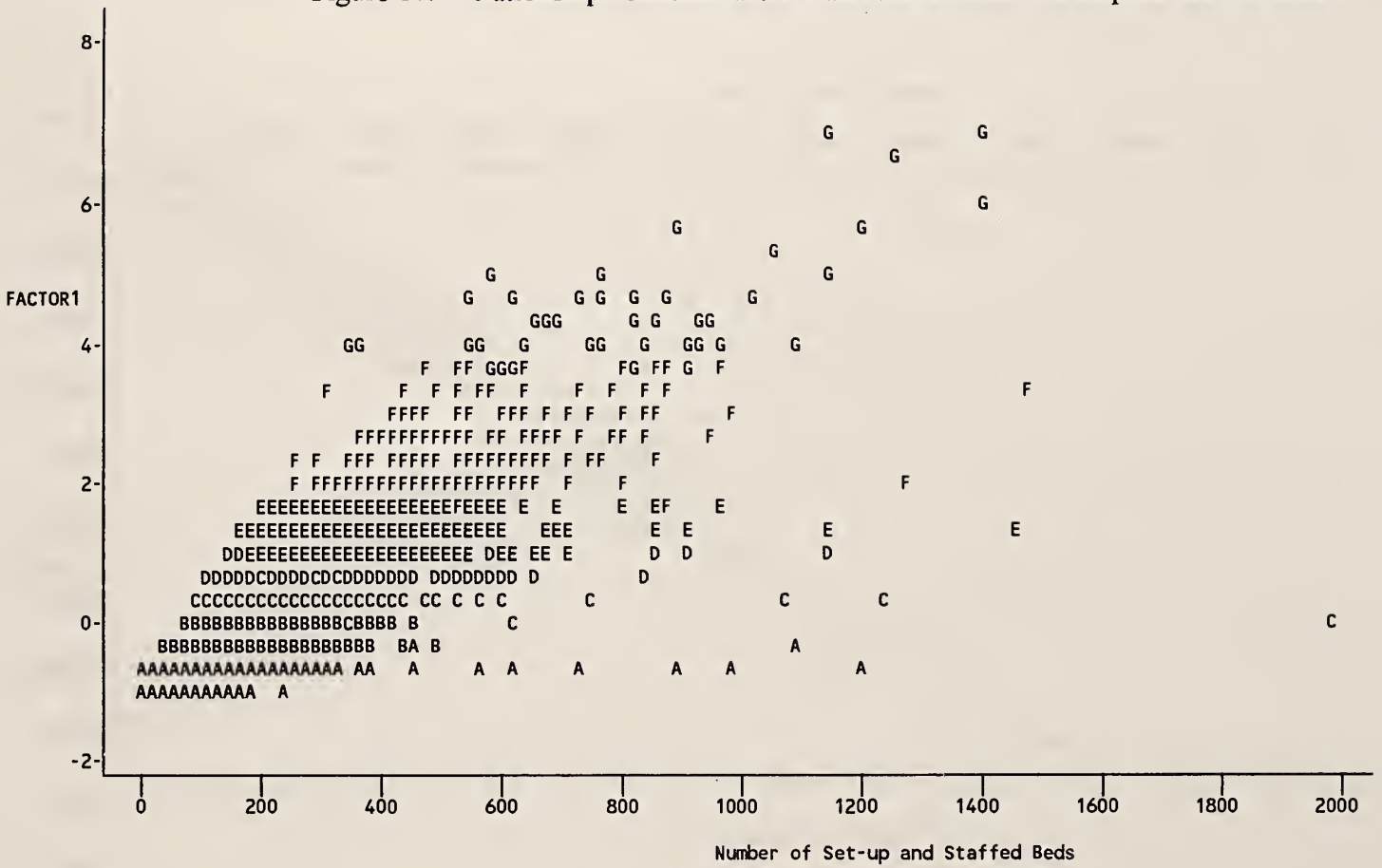
Alpha characters represent one or more hopitals in each of the 23 clusters. One cluster may superimpose another cluster.

Figure 9. Relationship between the total number of patient days and adjusted patient days.



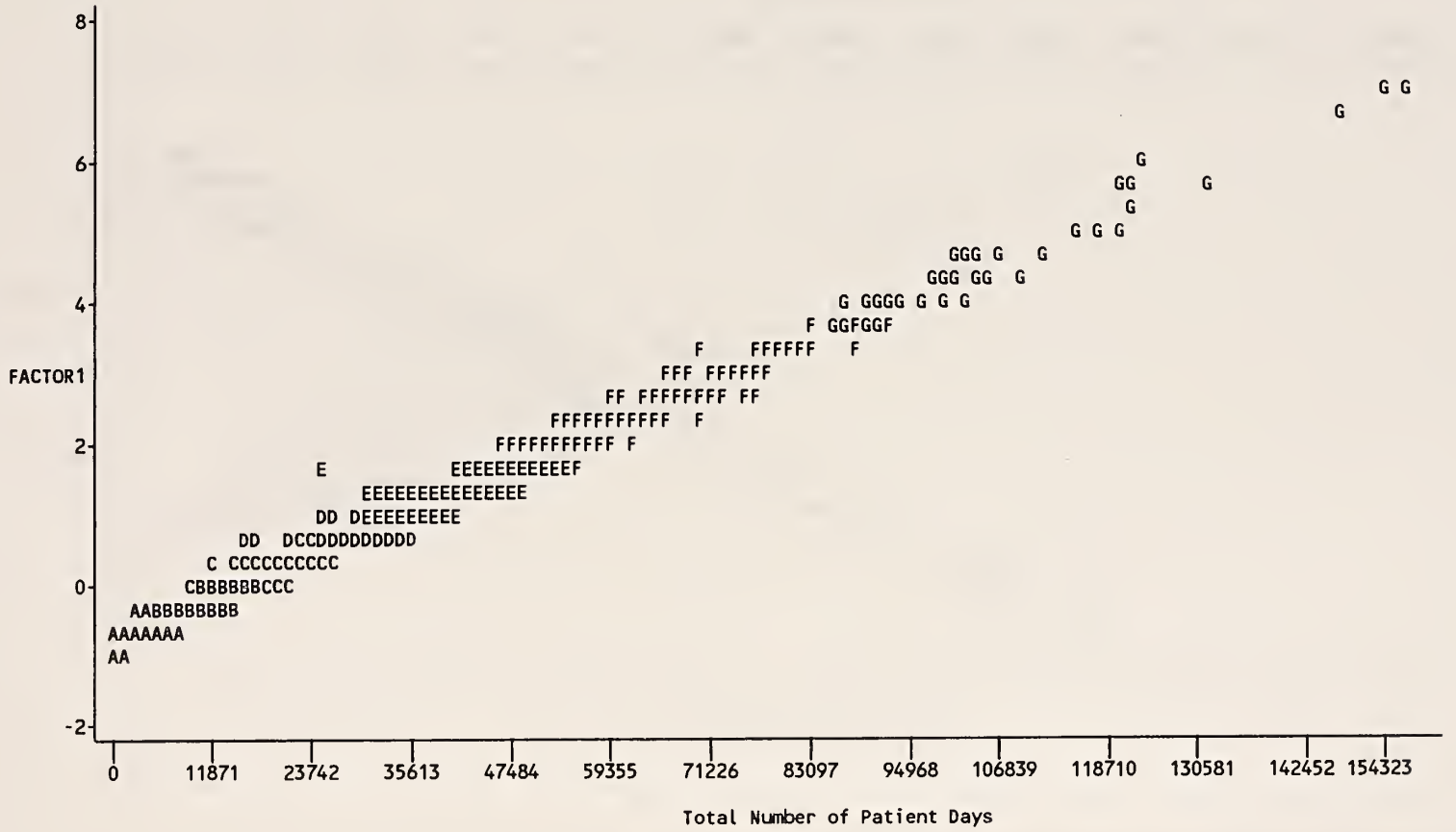
Alpha characters represent one or more hospitals in each of the 23 clusters. One cluster may superimpose another cluster.

Figure 10. Relationship between Factor 1 and the number of set-up and staffed beds.



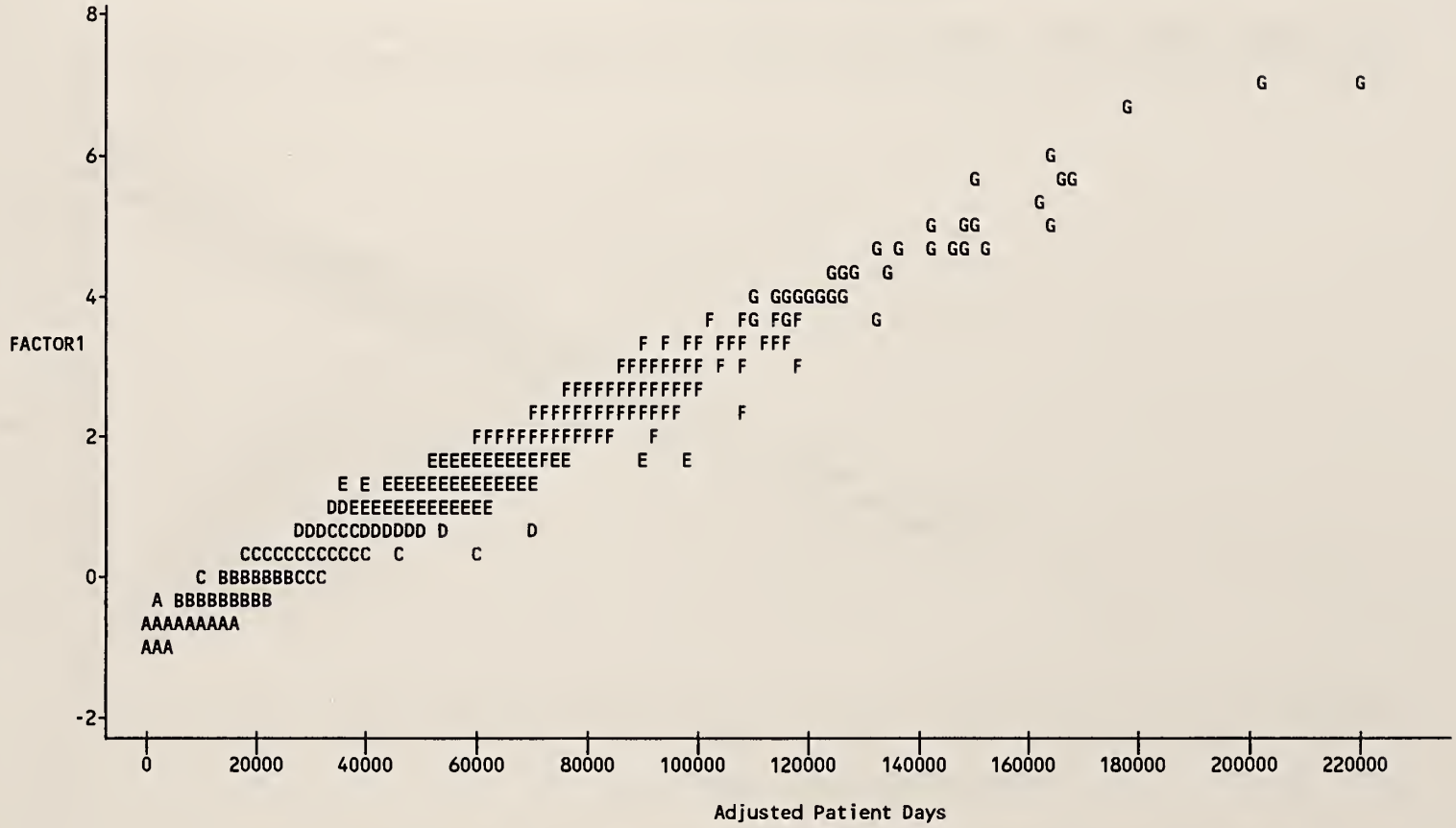
Alpha characters represent one or more hospitals in each of the 7 large clusters. One cluster may superimpose another cluster.

Figure 11. Relationship between Factor 1 and the number of patient days.



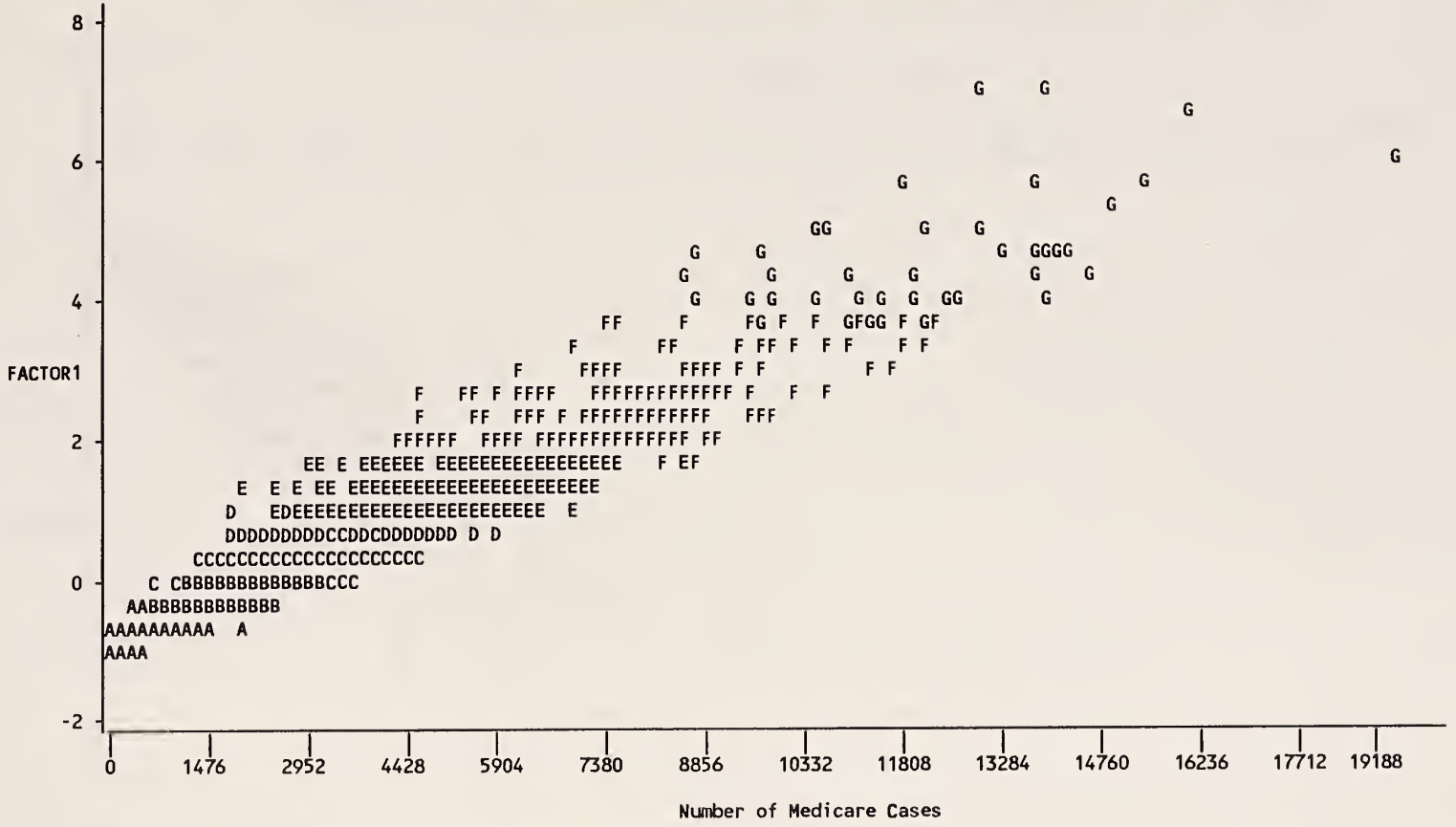
Alpha characters represent one or more hospitals in each of the 7 large clusters. One cluster may superimpose another cluster.

Figure 12. Relationship between Factor 1 and adjusted patient days.



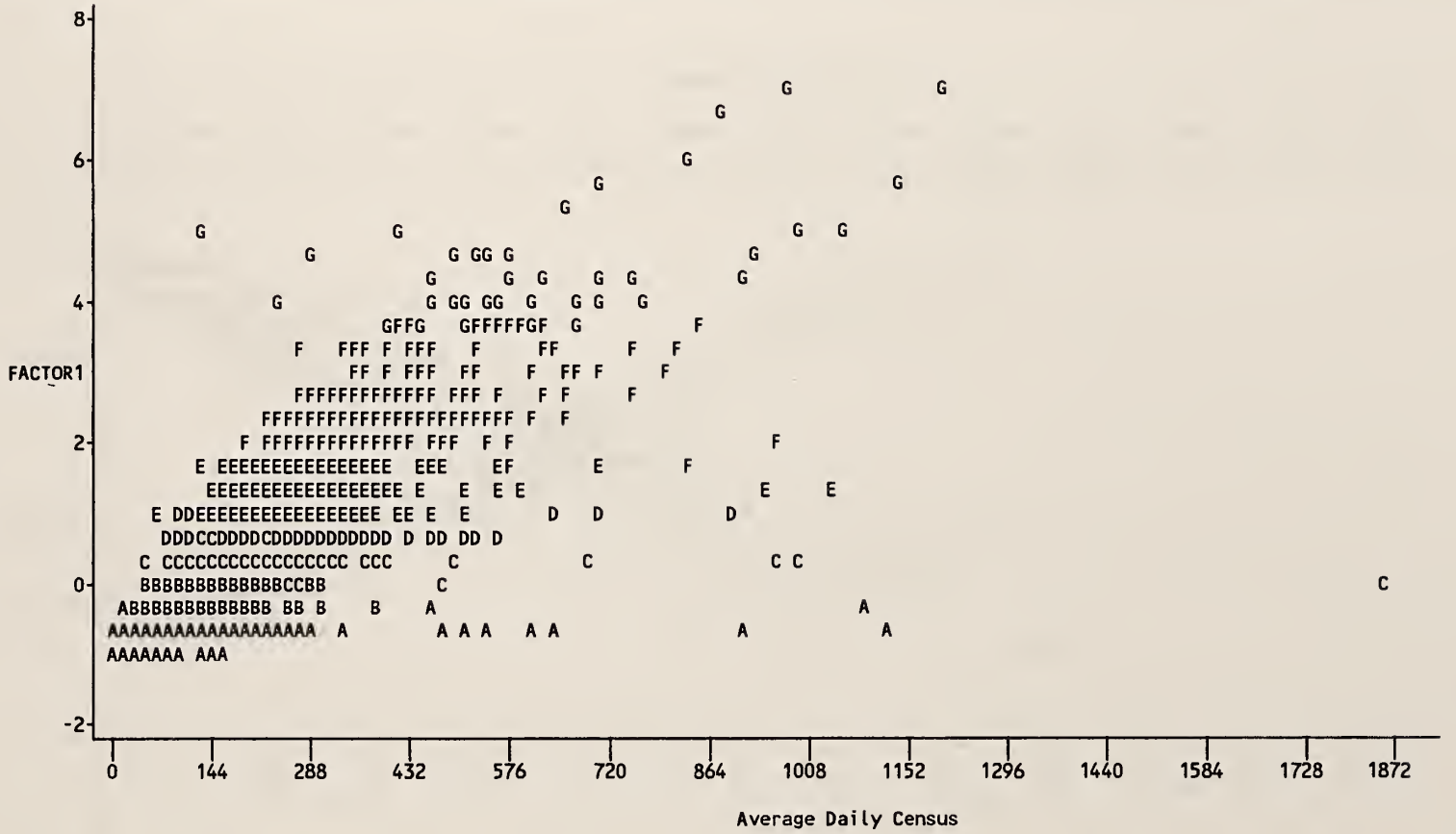
Alpha characters represent one or more hospitals in each of the 7 large clusters. One cluster may superimpose another cluster.

Figure 13. Relationship between Factor 1 and the number of Medicare cases.



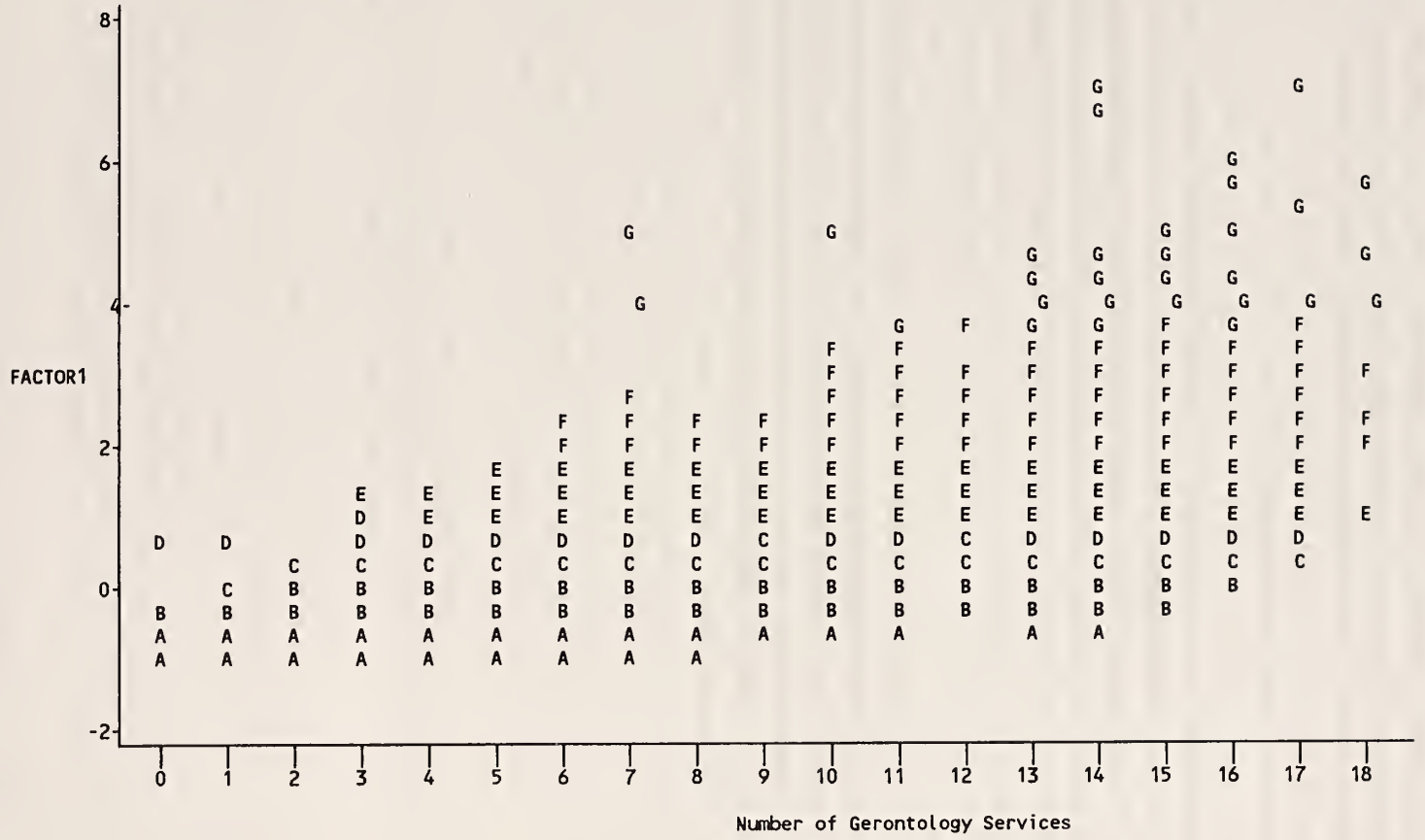
Alpha characters represent one or more hospitals in each of the 7 large clusters. One cluster may superimpose another cluster.

Figure 14. Relationship between Factor 1 and average daily census.



Alpha characters represent one or more hospitals in each of the 7 large cluseters. One cluster may superimpose another cluster.

Figure 15. Relationship between Factor 1 and the number of gerontology services.



Alpha characters represent one or more hospitals in each of the 7 large cluseters. One cluster may superimpose another cluster.

6.5. Face Validation

Seven PROs were requested to look at the second one-level cluster results. These PROs were for the States of Kansas (Kansas Foundation for Medical Care), Kentucky (Health Care Excel), Michigan (Michigan Peer Review Organization), New York (Island Peer Review Organization), Oklahoma (Oklahoma Foundation for Medical Quality), Tennessee (Mid-South Foundation for Medical Care), and Washington (PRO-West). Each PRO received a letter, Face Validation Forms for the 23 and 7 State Peer Groups, and a list of their state's hospitals shown in both the 23 and 7 hospital clusters (Appendix D).

Comments were received from all seven states. Several PROs noted that the 23 and 7 cluster results did not replicate pre-existing peer groups which were in their use at the time. These pre-existing peer groups included those created for HCFA's Cooperative Cardiovascular Project and those previously created and/or accepted by the PROs. They were often based on simple criteria such as urban vs. rural, or teaching vs. non-teaching. One PRO rejected the 23- and 7-cluster results because they did not match the PRO's previously, subjectively defined peer groups.

A major criticism was that specialty hospitals were included in the model when they should have been excluded. Since the DRG entropy variable in the model measures the diversity of DRGs for each hospital, it is a strong indicator of specialization. This fact was perhaps overlooked by the PROs. Besides, the design of this model was intended to include all hospitals. The second criticism related to the number of hospitals that were not included in the models because of the missing data.

7. Discussion

7.1. The Concept

This peer grouping project was conducted in an effort to develop a flexible, objective, and reproducible method for forming peer groups of acute care hospitals. Previous hospital peer groups used by most PROs and HCFA suffered in all of these three objectives.

First, since most of the peer grouping had relied on classifying hospitals with cut-off values based on the bed size or location, changes in the medical environment, such as merging hospitals, could make the old cut-offs obsolete. While obsolescence is inevitable to any peer-grouping method, such classification schemes are not readily adapted to a new environment since they rely on subjective judgements. By using a data-driven method for forming peer groups, as was done in this project, the peer groups can be reformed at regular intervals, allowing them to stay relevant over time.

Secondly, the classification methods based on one or two variables necessarily introduce subjectivity not only in the choice of the cut-off values, but also in the choice of the variables. To simplify a classification scheme, only a small number of variables could ever be used. Thus, simple classification schemes introduce subjectivity through the drastically limited initial choice of variables. Limiting classification variables to a small number also precludes building peer groups based upon many facets of a hospital's structure. By including many more variables and then using a factor analysis to reduce the dimensionality, such as was done in this project, a wide spectrum of hospital variables can be included. Using a large number of variables in the initial phase of forming the peer groups exploits as much information as is available. Using factor analysis to reduce the dimensionality compresses the large number of variables to a small number of characteristics. Thus, the method used in this project eliminates some of the subjectivity, makes the model more flexible to change, and includes more information, yet keeps the non-statistical explanation of the model simple.

Aside from classification methods, two peer-grouping methods using cluster analysis were developed by the PROs for Indiana and Michigan. The following concerns led the workgroup to reformulate the peer-grouping methods from anew.

A major concern was that of portability. Since this effort was intended to be easily transported, the data as well as the methods had to be portable. This was the primary reason for which neither of the two previous methods were used directly. The Indiana PRO peer grouping scheme required data from the Medicare Hospital Information Project. The Michigan extension of the Indiana method required some variables from the AHA Annual Survey of Hospitals which are no longer computed. Since the initial selection of variables is so important to the analysis, it was decided that the actual methods should be completely redesigned.

Instability of peer grouping was another reason for redesigning the methods. In the previous efforts, the pseudo statistics used to choose the number of peer groups were statistics which would perform best under the assumption of a mixture of multi-variate normal distributions.

Since hospital characteristics violate this assumption, there was no reason to assume that the pseudo statistics would perform well. Choosing the wrong number of peer groups, as these might, would misclassify the peer groups, and cause the actual peer grouping to be unstable. Such instability was experienced in the Indiana method of peer grouping -- when the initial peer groups were updated one year after they were initially formed, approximately 35 of the 120 hospitals changed peer groups. The Michigan peer grouping was never replicated on later data, so its stability was not known. Since the number of peer groups was based on the same pseudo statistics, it is possible that it would have been equally unstable. Experience had shown that there were some weaknesses with both the Indiana and Michigan models, and there was no evidence that these weaknesses would be avoided when moving from a state to a national level. Consequently, it was decided to try and develop a new clustering method.

7.2. Data

Accurate and current data are crucial to clustering hospitals into peer groups. The data used in these peer-grouping analyses were as accurate and current as were available to this workgroup. However, there are many areas where the data could and should be improved for future revisions and updates.

The data set used in this effort was originally intended to come entirely from HCFA. As stated above, the AHA hospital data were used rather than HCFA's poor quality POS file. Using the AHA data set had the advantage of accuracy, but it also had several disadvantages. First, the AHA data are proprietary, so the workgroup was constrained to use the 1994 data already owned by HCFA. Future use of the methods outlined in this report would require the extra expense of purchasing data if they would avoid this temporal constraint. The proprietary nature of the AHA data also limits its free exchange within the PRO community.

A second disadvantage of using the AHA data set is that engendered by the linking two data sets -- that of HCFA and that of the AHA -- whose sources differ. Multiple data sources necessitate the use of crosswalk files to link the data, such as the AHA crosswalk file used in this project. Crosswalks have an inherent problem in that their completeness lags behind that of the data which they link. This kind of problem was encountered in our analyses, as evidenced by the 410 providers whose entries were missing in the crosswalk. While many of the missing providers were merged by a primitive name and address matching scheme, and the matches were then manually checked for plausibility, it is likely that some mismatches could have occurred. These linking problems would be obviated if all of the data came from a single source.

Another disadvantage of using AHA data was that the data for many of the hospitals were missing either because they were not members of AHA or their response to the AHA surveys was incomplete. While the problem of non-response or incomplete response would not be solved by using HCFA data, the completeness of the hospital membership to the data set would be guaranteed. The estimated values of the data variables missing in the AHA files were extrapolated and tested at a regional level. As a result, a large number of previously excluded hospitals were successfully clustered into peer groups, but these additional hospitals had a

minimum impact on the composition of the peer groups. Alternatively, the variables with values missing for some hospitals could be excluded from the analysis. This would allow a greater number of hospitals to be eligible for clustering. With the exception of the number of gerontology services, which could be estimated from the case mix index, the exclusion of some other variables had very little impact on the final Factor which was used for clustering in the final model.

The choice of variables which were initially considered for the cluster analyses was motivated by the need to include variables which described the hospital environment, such as the size, teaching or non-teaching, for-profit or non-profit, urban or rural location, out-patient activity, diversity of services, staffing levels, efficiency in terms of staffing levels, and accreditation status. These variables could not only describe the current hospital environment, but would be useful for future endeavors. Examples of this were the choice of organizational variables, such as membership in hospital chains or HMOs, and status as a community, non-profit, or for profit hospital. These variables were added to the usual size variables, since the level of HMO penetration varied greatly across the nation. Furthermore, approximately 700 not-for-profit hospitals per year have been bought by for-profit hospitals since 1994.[24] Since differing care in these hospitals is debatable,[25-27] and to make the model flexible for future changes, these variables were also selected initially. Lastly, the average daily census was included in the initial choice of variables to adjust for the increasing proportion of procedures which were performed in the outpatient setting. The goal was to select a comprehensive set of variables which measure the most important aspects of hospital structure and services that impact the present and future Medicare populations.

There are some factors that make the current health care environment unique compared with previous years. For example, the hospital market has been greatly affected by managed care and the level of impact varies greatly by geographic region. A recent study indicated that as the Health Maintenance Organization (HMO) penetration in a region increases by 10 percent, the number of hospitals or hospital beds would decrease by 5 percent.[28] With the medical technology improvements, more health care services are available at the outpatient setting , and as a result, the hospital occupancy rate has been decreasing in recent years.[29] Hence the average daily census (ADC) is regarded as a better proxy for hospital size than the number of hospital beds. It is consistent with the fact that HCFA recently redefined the graduate medical education adjustment formula based on ADC.[8]

It is likely that hospital practices are influenced by the current health care trends, and they may relate to hospital performance and health care outcomes. Hence, it was important to reflect these new changes in the hospital peer grouping. However, researchers are cautioned by AHA that there may be considerable variations in the economic climate, demographic characteristics, personal factors, and health care financing and payment policies among hospitals.[19]

7.3. Methodology

One of the innovative features of this work is its use of bootstrapping to objectively determine the proper number of peer groups using the structure of the data. The use of bootstrapping

minimizes the problems inherent in the visual interpretations of dendograms or the use of statistics based upon data which follow a multi-variate normal distribution. The downside of bootstrapping was the extremely large amount of time it required to run. Much of this time consumption can be directly traced to the difficulty encountered in coding bootstrapping into SAS. In order for future implementations to be practical, a more efficient algorithm would be needed for forming the bootstrapping samples.

7.4. Statistical Analyses

The development of the desired peer grouping model required a process of trial and error. In this process, four models were tested -- two one-level and two two-level models. In each model, a factor analysis of selected variables followed by a cluster analysis were conducted. Since the first one-level model was the most basic and parsimonious model, the results were entirely determined by the mainstream statistics. This model was rejected because of its poor face validity. Peer groups appeared to be clusters based solely on the size of the hospital. The influence of other variables, such as the diversity and complexity of services, were negligible.

The problem of a single variable washing out the rest is a common problem in factor analysis. The suggested work-around was a two-stage model -- first, stratifying the data on the most influential variables (in this case, size), and then running factor and cluster analyses on each stratum. Several variables used in these analyses measured size, for example, the number of set-up and staffed beds, total patient days, adjusted patient days, number of Medicare discharges, and average daily census. These findings lead to testing the two-level models.

Total patient days and MSA population size were used in the first two-level model's cluster analysis which resulted in three strata according to the hospital size -- small, medium, and large. Principal factor and cluster analyses were then run on each stratum. This model was also rejected after face validation, because the peer groups did not make sense.

The second two-level model was considered because other variables indicating the size of the hospital could have been better predictors. Here, the total number of patient days and the number of set-up and staffed beds were used in the cluster analysis. Principal factor and cluster analyses were run on each hospital size stratum. The results showed a better stratification of hospitals, and this appeared to be the best model up to this point. The hospital peer groups discriminated on both the size and complexity. The face-validity tests, however, uncovered several problems. There were many hospitals which did not belong to the peer group to which they were assigned; there were too many peer groups in the final solution; and the peer groups were ill-defined. Hospitals in the same peer group were very diverse in terms of the size, diversity of services, staffing, etc. Consequently, it was concluded that this model did not have sufficient practical applicability.

Instead of the usual least squares factor analysis, the maximum likelihood factor analysis was tried. The advantages of the latter technique include more precise solution which would be less likely influenced by extreme scores and would not require an assumption of an underlying multi-variate normal distribution. The disadvantages of using this method were the added

computational expense of iterative calculation and potential difficulty in arriving at a final solution. The results of the second one-level model using the maximum likelihood factor analysis followed by a cluster analysis showed one dominant factor which explained 99 percent of the variation in all of the input variables. The face validity tests suggested that this model would be acceptable. Furthermore, it had the advantage of being parsimonious, that is, it was not determined by a pre-selected set of variables.

The 23 clusters identified in this model were quite distinct, well defined, and valid as determined by the face-validity assessments. However, it would have the disadvantage of assigning very small numbers of hospitals to some of the clusters (peer groups) within a small sample, such as acute care hospitals in a small state. A 7-cluster model was, therefore, developed by running a final cluster analysis on the 23 clusters. A comparison of the two models suggested that the 23 clusters were in effect aggregated into seven clusters (Table 15). Such an aggregation loses some precision, definition, and clear discrimination, but it overcomes the problem of small numbers. The users of these peer grouping models can choose the one that suits their local requirements.

7.5. Comments from PROs and HCFA Staff

The workgroup presented the preliminary results of this project at the American Health Quality Association's (AHQA) Technical Conference and at HCFA's Grand Rounds in Baltimore, Maryland in 1997. Comments followed by discussion are described below.

7.5.1. AHQA Conference

In response to questions, the workgroup explained that SAS assigns numbers to peer groups as they are identified from the data and, therefore, does not order them along a continuum. These random numbers were reviewed and peer group numbers were assigned using the Factor scores and an ordinal scale. Because of the ordinal numbers, users of these peer grouping models will be able to recluster the peer groups within an individual state or region by combining peer groups which are contiguous in the ordinal numbering. Indeed, the factor score for each hospital could be used to recluster the peer groups hospital by hospital, if needed.

For ease of manipulation and use, two different sets of peer groups will be available. The first set contains 23 peer groups in a continuum. The second set contains only 7 peer groups. Both models have ordinal numbers assigned. Depending upon the need of the user, one or the other of these peer groupings will be suitable.

It should be remembered that the 23-cluster model was designed using data from as many hospitals across the United States as possible. Therefore, the results are the best grouping of the nation's acute care hospitals. This method may be less satisfactory in peer grouping a single state's or region's hospitals. Since the model first assigns the Factor scores and then creates clusters of peer hospitals, a decrease in the number and variation within any group, state, or region will affect the Factor scores and the resulting peer groups. The peer groups created by this model would be most suitable for national comparisons and work best at that level.

Using the computer program developed in this project, an individual state or region could produce new hospital peer groups. This activity would be time consuming, but would produce a unique solution for the hospitals entered into the model. Considerable information was provided at the conference and additional information will be available in the user's guide. It was realized that one of the best use of this peer grouping model would be in case-control studies to evaluate the effectiveness of quality improvement interventions of HCQIP projects.

7.5.2. HCFA Grand Rounds

Among the questions asked by staff at HCFA's central and regional offices were questions concerning the clustering methodology, dendograms in particular. Since bootstrapping was used to determine the best number of clusters, subjective decision of where the dendogram appeared to suggest the best number of clusters was not necessary. In addition, a dendogram assumes that the algorithm used was hierarchical; and since both Wards and Average Linkage algorithms were used, the model suffered from chaining problems which precluded the use of a dendogram.

A technical question revolved around the use of standardized variables, non-linear transformations of the variables, and other clustering methodologies. Since the variables had been standardized, and hence normalized, no transformation of the variables was used. Neither the Leo Brindman classification nor regression tree methodology was used to produce the model. No matter which of the possible clustering methodologies is chosen, a great deal of subjectivity exists.

The peer groups resulting from this model are superior to the crude peer grouping based on hospital size, region, and/or urban vs. rural dichotomy, because they are at the national level. The other advantage is that "true" peers of a hospital in other states can be easily identified. This is particularly true of the largest hospitals in many states which actually have no true peers in those states. This model also allows for differentiation among the largest hospitals within large, urban states. In addition, several of the variables related to efficiency had never been used in previous hospital peer grouping efforts. During face validation of the peer groups, any individual hospital that seemed to be misallocated was actually placed in the correct peer group upon examination of its characteristics and those of its peers. This indicated that the cluster analysis was more accurate than the intuitive placement from previous observation and experience.

HCFA staff recommended that the peer groups be highly publicized within the PRO community and be presented at research conferences and symposia, because the value of the peer groups would become more and more obvious if they were integrated by the PROs in HCQIP projects and refined through use. It was also realized that the peer groups would be useful in policy and health service utilization discussions.

This peer grouping methodology would also be useful to hospitals and PROs in targeting the appropriate audience for health care quality improvement, putting similar institutions in touch with each other for networking purposes, providing benchmark information for an institution

who, while similar, is not an immediate geographic competitor, providing hospitals with a source for process improvement information that is immediately applicable to their institution.

In the context of HCQIP, PROs have a very important role in describing the hospital peer grouping methodology used in their projects so that the hospitals and their trade associations can become comfortable with the peer groups. Only when the hospitals understand that their comparative group is really made up of their peers, will they accept the comparisons and demonstration of opportunities for improvement made by the PROs and HCFA.

These peer groups may not be applicable for specific clinical topics such as hospitalization for cancer or cardiovascular diseases. Consequently, hospitals with a particular clinical area of expertise may not be clustered in one peer group. The peer grouping program can be run on a set of variables which are relevant to a specific clinical topic. The resulting peer groups will provide the same level of comparability for that clinical topic that these peer groups do for the general Medicare patient population. Careful thought will be required to select the variables relevant to the hospital(s)' characteristics, such as services and staffing, and the patients' characteristics. A computer program for factor and cluster analyses, and a user's guide (Appendix F) will be provided to all PROs and other interested users.

7.6. Face Validation

Staff at the seven PROs who validated the two clustering models also provided useful comments. The age of the data used in the cluster analyses was perceived as a critical problem. The health care industry in the United States may have changed radically since 1994, particularly the changes in the organization of hospitals and health care systems. Apparently, several hospitals which were independent facilities in 1994 had been affiliated with or even purchased by other facilities prior to 1997, the time of validation. However, for many hospitals, the 1994 data are still valid representation of their characteristics and population; for others the current peer groups no longer define true comparable facilities.

The 23-cluster model may not be suitable for many states with small number of hospitals. It would be difficult to maintain the confidentiality of hospitals if individual peer groups had very few members. For the states with large rural populations and small rural hospitals, as many as 50 percent of the hospitals were found within the first peer group. Subgrouping of these hospitals would be very helpful for the PROs in those states. On the other hand the 7-cluster model was perceived to be less problematic. Several PROs found that the seven peer groups approximately matched their pre-existing peer groups based on geographic location, relative hospital size, and urban-rural definitions.

The problems of data confidentiality in the 23-cluster model and the problems related to a lack of definition in the 7-cluster model can be resolved by the application of the software program which will be provided to the users. For example, the ordinal numbers can be used to recluster the peer groups. The user will be able to define the desired number of peer groups that would provide the greatest flexibility and validity in a given situation.

During the validation process, the PRO for Oklahoma found that non-acute and transitional hospitals were clustered in the same peer groups as acute care hospitals. This was obviously a function of the hospital-specific Factor scores which were based on the data provided by the hospitals in the 1994 AHA survey. Changes in the characteristics of these hospitals since 1994 resulted in dissimilarity between the two types of hospitals seen in 1997. It was also noted that the hospitals which provide care to the Native American population in Oklahoma did not cluster into a single peer group, although they appear to be a distinct group of hospitals. This lack of specificity was obviously due to the fact that patients' demographic characteristics, such as race, were not used as variables in the cluster analysis. These data were not collected by the AHA surveys, and there was no evidence in the literature regarding the importance of patients' race in peer grouping acute care hospitals.

The model was not as discriminating as had been hoped in the identification of specialty hospitals. Entropy, a measure of the diversity of DRGs within a hospital's patient population, was the variable used to identify specialty hospitals. This measure did not result in the level of discrimination as had been hoped, since specialty hospitals had been clustered within the same peer groups as the general hospitals. Inclusion of speciality hospitals was also criticized by some of the PROs who validated the results (See Section 6.5). A line has been added to the computer program for clustering which will omit the speciality hospitals if peer grouping of only general hospitals is desired.

The state-specific definition of "tertiary" also appeared to present some difficulty to the PROs who conducted face validation. A tertiary care center appears to be relative to the size of the state's population. In other words, a tertiary care hospital in a rural state may not be of the same size and complexity as a tertiary care hospital in a large urban state although both would be seen as tertiary care centers in their respective states. So too does the definition of a community hospital appear to differ depending largely upon the urbanization of the state's population. A community hospital is seen as being different from a tertiary referral center, no matter what the size or complexity of the tertiary hospital. These subjective assessments and definitions specific to individual organizations are to be expected. Hence, the main purpose of using statistical analyses for peer grouping was to avoid subjectivity.

7.7. Limitations

Eight hundred and eighty-two hospitals (16 % of all hospitals) were excluded from the analyses because of missing information. While some large and prestigious hospitals were excluded because of missing data, some university hospitals did not participate in AHA surveys because they were not members of AHA. These exclusions could create some degree of selection bias in peer grouping hospitals, particularly in small samples such as hospitals in a state. To avoid loss of hospitals from the cluster analysis and/or selection bias, the average values of the variables in the data set can substitute for the missing values. Using HCFA data in regression equations, it was possible to classify hospitals which were initially excluded. Although this technique was not used in this peer grouping model, it is included in the computer software as an option for replicating hospital peer grouping which will be available to PROs and other users.

The goal of the cluster analyses was to classify all hospitals using available HCFA and/or AHA data. Specialty hospitals, such as hospitals which specialize in weight loss, cosmetic surgery, and orthopedic surgery, were not excluded. When examining the preliminary data, extreme values were found for all of the variables. This may explain why as many as 23 clusters were needed to classify the hospitals. Other users of the peer grouping model may wish to trim outliers and re-run the programs. If this analysis is repeated on the same data, the decision to include all hospitals should be revisited.

Data collected in 1994 were used in the cluster analyses. Given the fluid hospital market, this peer grouping model may have limited use. In order to update this model, users or HCFA will have to collect recent AHA survey data and revise the model.

Cluster analyses using cross-sectional data from one survey may not correctly portray hospital peer grouping over time. Replication with current data will be necessary to assess the stability of the peer groups. It is also possible to measure changing market factors with trended data by comparing the statistics for each variable over time. For example, peer grouping in the future may be dependent upon HMO/PPO and health system memberships instead of size and FTEs. Similarly, financial data were not used in the analyses, and it may be that current financial data as well as other AHA variables are highly predictive. Users of this methodology can add variables to the models. Some financial data as well as information on other hospital characteristics are available from the AHA surveys.

This peer grouping model may not be acceptable to some hospitals who fail to recognize that some of their peers are their competitors, since competitor hospitals are likely to be of similar size and in the same region. Clear understanding of the concept of objective peer grouping based on 20 variables for most of the hospitals in the country should satisfy their concerns and promote acceptance.

This 23-cluster model may not be appropriate in certain situations. For example, this model will not provide correct peer grouping if the hospitals are of a certain type, in a specific region, or offer a selected service, such as open heart surgery, transplantation, or chemotherapy for cancer. Future users of this model can use the programs and methodology if hospitals are selected based on their geographical location, teaching status, size, types of services, or other criteria.

In the course of validating the 7- and 23-cluster models in their states, several PROs found a few hospitals, usually less than 10, which were in the wrong peer groups. This could be due to recent changes in the market, or flawed data, or the limitation of the model. This model should be extensively validated using recent data. It may be that this clustering model is not appropriate for certain groups of hospitals, for example, areas with high managed care concentration, or hospitals with selected services.

7.8. Uses

During 1998, a few PROs requested lists of acute care hospitals in their states peer grouped into

seven clusters and into 23 clusters. The reported intent was to compare their "crude" peer grouping method with these models and/or use the seven or 23 cluster model to compare hospitals within their states. These comparisons were conducted in an attempt to compare the quality of health care provided by a group of hospitals with that of their peers.

In 1997, HCFA's Regional Office in Kansas City used the newly developed hospital peer groups to evaluate quality improvement projects on the topic of radical prostatectomy conducted by PROs in ten states between 1992 and 1996. Historical control hospitals were identified using several matching criteria which included the peer status. The study successfully evaluated the effectiveness of the PRO project interventions by comparing reductions in the rates of radical prostatectomy procedure in the intervention hospitals with those in the control hospitals.[30]

Peer grouping for defining control hospitals can be used in most situations in HCFA's Health Care Quality Improvement Program and other studies in which individual or groups of hospitals are prospectively or retrospectively compared with each other. It would also be quite reasonable to use this methodology for comparing favorable or adverse outcomes of care provided by hospitals.

The variables selected for peer grouping hospitals were used to create factors and hospital clusters which in turn characterized the current hospital environment. In the future, changes in the health care industry could be measured by examining changes in the composition of these factors and clusters.

8. Recommendations

At the completion of the work of peer grouping acute care hospitals which participated in Medicare in 1994, the workgroup has several recommendations. These recommendations are for HCFA, PROs, and hospital trade associations.

8.1. Recommendations to HCFA

8.1.1. The development of these hospital peer groups has been an arduous task, and one that inherently contains a considerable amount of subjective judgement. As with all research, the results need to be tested to ascertain their robustness when applied to the needs of the users. It is, therefore, recommended that HCFA uses the peer grouping model in four specific ways as quickly and as often as possible:

- Designing national and regional health care quality improvement projects;
- marketing national and regional projects to hospitals by presenting the baseline and benchmark data to demonstrate opportunities for improvement;
- evaluating quality improvement projects within and across states by analyzing comparable hospitals' experiences and performance in quasi case-control studies; and
- applying the peer groups to HCFA's own policy and operational research.

8.1.2. Some of the comments from the PROs who face-validated the two cluster models were related to the age of the data used in the analyses. The workgroup was aware of these limitations of the data, and recommends that HCFA rerun the cluster analyses at regular intervals to accommodate the changes in the health care industry and changes over time. Updated data for all of the variables used would provide the opportunity to examine the trends over time and to retest the reproducibility of the model.

8.1.3. Since many of the variables used in clustering describe hospital characteristics, a consistent source for these data is necessary. The workgroup decided to use hospital data from the American Hospital Association's (AHA) Annual Survey of Hospitals in preference over the HCFA Provider of Service (POS) file (See Section 4.1). The workgroup has two recommendations in the area of the hospital characteristics' databases:

- The POS file was found to be out-of-date, incomplete and inaccurate. Therefore, the workgroup recommends that HCFA fully test the accuracy of the POS file, develop a process to improve it, and assure its viability. In this way, up-to-date and accurate hospital data would be available directly from HCFA data files, and the PROs and other users would not be dependent upon AHA for the data; and
- many hospitals had to be dropped from the database because the AHA-HCFA crosswalk was either too old or too inaccurate to correctly identify them. Therefore, the

workgroup recommends that HCFA and AHA work to update and improve the current provider crosswalk program.

8.1.4. The workgroup also feels that HCFA should take an active role in continuing the work on peer grouping of hospitals. The recommendations in this area include having HCFA sponsor other research, act as a clearinghouse for information received from other developers and users, and provide a forum for "lessons learned" by others. Specifically, all of these activities would encourage discussion of the appropriate and innovative uses of the peer groups, the effectiveness of the current peer groups, and state and regional level applications of the models.

8.2. Recommendations to PROs

8.2.1. As with all research, PROs should not use the data or the clustering blindly. PROs should examine the data and the results of clustering before applying these results to current health care quality improvement or research projects.

8.2.2. The workgroup strongly recommends that PROs routinely use the peer groups in their HCQIP efforts and provide feedback about the results to HCFA's clearinghouse. PROs have been advised by HCFA regarding the strategy for data collection for national health care quality improvement projects to be used in the 6th Scope of Work (SoW) for PROs beginning in 1999. HCFA will develop measures of the quality indicators in these national projects. However, the state-specific sample size of these data will not allow estimates at the individual health care provider level. Consequently, the quality indicator measures will have to be computed for groups of providers. If the opportunities for improvement are to be demonstrated at the individual hospital level, the hospitals in a group would have to be comparable to each other so that the average measure would be representative of the measures for individual hospitals. Peer grouping of hospitals within a state would be a convenient way of grouping hospitals with similar characteristics and expectations.

8.2.3. PROs should use peer grouping in presenting the data to providers and demonstrating opportunities for improvement. By briefly explaining the peer grouping model and convincing the providers that they are indeed being compared with their peers, PROs can assure the reliability of their data. This will facilitate collaboration.

8.2.4. PROs should use the peer grouping model and evaluate the impact of their projects by comparing improvements attributable to their project interventions with those changes achieved by peer providers who did not collaborate in the project. Furthermore, according to the 6th SoW, PROs will be required to conduct a local quality improvement project which will target a predefined disadvantaged group of Medicare beneficiaries in an attempt to reduce disparities between health care received by the disadvantaged beneficiaries and all other beneficiaries living in the state. Each of these projects shall have at least one intervention group and a control or comparison group. If a standard experimental design for these projects is not feasible, PROs should consider using peers of the intervention hospitals as their controls.

8.2.5. The workgroup recommends that each PRO establish a research team to work with HCFA to rerun and improve the current peer grouping model. Using the programs provided by HCFA, the PROs may rerun the clustering to monitor intra-state changes over time and provide this additional information to HCFA's clearinghouse.

8.3. Recommendations to Hospitals and Their Trade Associations

8.3.1. As with all other data driven activities, the results of the analyses conducted by the workgroup are highly dependent upon the accuracy of the data provided to HCFA and to AHA. It is, therefore, recommended that the hospital trade associations work to continuously improve the hospital and patient data that are submitted to AHA and to HCFA.

8.3.2. The workgroup recommends that the hospital trade associations use these peer groups for internal research and policy development. Hospitals may also find this peer grouping methodology to be a convenient tool for evaluating their data.

8.3.3. Hospitals and hospital trade associations are further encouraged to use peer grouping in their outcomes and other research, and in routine monitoring of their performance and continuous quality improvement activities.

9. **Glossary:** List of Acronyms

ADC - Average daily census
ADJPAT - Adjusted patient days
AHA - American Hospital Association
AHQA - American Health Quality Association (formerly AMPRA)
AMPRA - American Peer Review Association

BDTOT - Number of set-up and staffed beds

CASES - Total Medicare patient discharges
CMI - Case Mix Index

DRG - Diagnosis Related Group

FTE - Full Time Employee
FTEH - Total hospital FTE's
FTERN - Total RN FTE's

HCFA - Health Care Financing Administration
HCQIP - Health Care Quality Improvement Program
HMO - Health Maintenance Organization

MHS - Membership in a health system
MPN - Medicare Provider Number
MQIS - Medicare Quality Indicator System

ORHADSUR - Percentage of Medicare discharges for which at least one procedure was coded

POS - Provider of Services Database
PPS - Prospective Payment System
PRO - Peer Review Organization

RESADC - Total residents/Average daily census

SAS - Statistical Analysis System
SDPS - Standard Data Processing System
SNF - Standard Analytical File
SOW - Scope of Work for PROs

TOTLOS - Total length of stay/patient days
TRANS - Percentage of Medicare discharges transferred out of the hospital

UPIN - Unique Physician Identification Number

10. References

1. Jencks SF, Wilensky GR. The health care quality improvement initiative: A new approach to quality assurance in Medicare. *JAMA* 1992;268:900-903.

2. Sheikh K. Defining and achieving quality of medical care. *Am J Medical Quality* 1998;13: 59-62.

3. Jencks SF. Measuring quality of care under Medicare and Medicaid. *Health Care Financing Review* 1995;16:39-54.

4. Phillip PJ, Iyer RN. Classification of community hospitals. *Health Services Research* 1975; 10(4):349-68.

5. Elayet HA. Entropy in the hierarchical cluster analysis of hospitals. *Health Services Research* 1978; 13(4):395-403.

6. Klastorin TD. Determination of alternative hospital classifications. *Health Services Research* 1981;16(2):205-19.

7. Stefos T, LaVallee N, Holden F. Fairness in prospective payment: a clustering approach. *Health Services Research* 1992;27(2):239-61.

8. Morey RC. Estimating the hospital-wide cost differentials warranted for teaching hospitals. An alternative to regression approaches. *Medical Care* 1995;33(5):531-52.

9. Bodenheimer TS, Grumbach K. Understanding health policy: Clinical approach. Norwalk, CT: Appleton and Lange, 1995.

10. Iglehart JK. Early experience with prospective payment of hospitals. *New England Journal of Medicine* 1986:314(22):1461.

11. Lee KY. Classification of hospitals based on two-stage multivariate methods. *Clinical Performance and Quality Health Care* 1995;2(4):195-9.

12. Zodet MW, Clark JD. Creation of hospital peer groups. *Clinical Performance and Quality Health Care* 1996;4(1):51-7.

13. Klastorin TD, Watts CA. A current reappraisal of Berry's hospital typology. *Medical Care* 1982;20(5):441-9.

14. Alexander JA, Evashwick CJ, Rundall T. Hospitals and the provision of care to the aged: a cluster analysis. *Inquiry* 1984;21(4):303-14.

15. Dilts D, Khamalah J, Plotkin A. Using cluster analysis for medical resource decision making. *Medical Decision Making* 1995;15(4):333-47.

16. SAS Institute Inc. SAS/STAT User's Guide; Vol. 1, Ver. 6, 4th Ed. Cary, NC: The SAS Institute, 1990.

17. Health Care Financing Administration. Cooperative Cardiovascular Project Tutorial. Baltimore, MD: The Health Care Financing Administration, 1995.

18. Tonn S, Ellen P. Factor analyzing provider characteristics: Data sources, methods and applied uses. St. Petersburg Beach, Florida: Quality Improvement Organizations Tri-Regional Conference, 1996 (Presentation).

19. American Hospital Association. Hospital stat: Emerging trends in hospitals. Chicago: The American Hospital Association, 1996

20. Dillon WR, Goldstein M. Multivariate analysis: Methods and Applications. New York: John Wiley & Sons, 1984.

21. Norman GR, Streiner DL. *PDQ statistics*, Toronto: B. C. Decker Inc., 1986.

22. Efron B. The jackknife, the bootstrap, and other resampling plans. Philadelphia: Society for Industrial and Applied Mathematics, 1982.

23. Jain AK, Moreau JV, Bootstrap technique in cluster analysis. *Pattern Recognition* 1987;20(5):547-68.

24. Editorial. Who protects public as hospitals aim for profits? Arlington, VA: USA Today, Dec. 17, 1996.

25. Ozcan YA, Luke RD, Haksever C. Ownership and organizational performance. A comparison of technical efficiency across hospital types. *Medical Care* 1992;30(9):781-94.

26. Kuhn EM, Hartz AJ, Krakauer H, Bailey RC, Rimm AA. The relationship of hospital ownership and teaching status to 30- and 180-day adjusted mortality rates. *Medical Care* 1994; 32(11):1098-108.

27. Geller SE, Burns LR, Brailer DJ. The impact of nonclinical factors on practice variations: the case of hysterectomies. *Health Services Research* 1996;30(6):729-50.

28. Chernew M. The impact of non-IPA HMOs on the number of hospitals and hospital capacity. *Inquiry* 1995:32(2):143-54.

29. Feldstein PJ. Health Care Economics. 4th Ed., Albany, NY: Delmar Publishers Inc., 1993.

30. Sheikh K, Bullock C. Evaluation of quality improvement interventions: A non-experimental model. Austin, Texas: American Health Quality Association Technical Conference, 1998 (Presentation).

Appendix A

Summary of articles on hospital peer grouping in which peer groups were created using statitical methods

Title	Classification of Community Hospitals	Entropy in the Hierarchical Cluster Analysis of Hospitals	Determination of Alternative Hospital Classifications
Authors	Phillip, PJ and Iyer, RN	Elayat, HA et al	Klastorin, TD and Watts, CA
Journal	Health Services Research (1975)	Health Services Research (1978)	Health Services Research (1981)
Primary Objective	To provide an alternative hospital classification system for the use of Social Security Administration	To determine the number of clusters using the concept of entropy	To present a stabe hospital classification method when constraints on the classification process are recognized
GroupingVariables	Product (beds, hosp. type, # advanced services, long-term care, pay, assets, surgery), External (MSA, pop, income, physician, gender, age)	Cost (# services, # advanced services, long-term care, surgery, beds, assets, nurses, non-medical personnel, income, physicians)	Cost (wage, # services, birth per patient, outpatient visits per patient, surgeries per patient, income, age distribution)
Number of Hospitals	4,550 National Hospitals	218 Michigan hospitals	200 random sample hospitals
Grouping Method	Stratified Factor Analysis and Cluster Analysis	Hierarchical Cluster Analysis	Principal Component Analysis & Hierarchical Cluster Analysis
Number of Peer Groups	71 groups (64 hospitals per group on the average)	16 groups based on entropy with respect to cost per day	32 groups (avg 6.25 hospitals) using expected distintiveness
Post-Evaluation	ANOVA, Discriminant Analysis	None	None
Application of Peer Groups/Method	Rate review, self-evaluation, experimental design	Reimbursement system	Reimbursement system
Comments	Stratum-specific peer groups (five strata by #beds), 484 nonclusterable isolates	-	Cluster variables were regression coefficients from avg. expense per case and factor scores

Appendix A

Summary of articles on hospital peer grouping in which peer groups were created using statical methods (continued)

Title	A Current Reappraisal of Berry's Hospital Typology	Hospital and the Provision of Care to the Aged	Fairness in Prosepective Payment: A Clustering Approach
Authors	Klastorin, TD and Watts, CA	Alexander, JA, et al	Stefos, T et al
Journal	Medical Care (1982)	Inquiry (1984)	Health Services Research (1992)
Primary Objective	To evaluated Barry's five hospital types using more up-to-date hospital data	To classify hospitals based on geriatric services provided and to identify hospital characteristics providing such services	To cluster hospitals to establish competitive prosepective payment system for VA hospitals
GroupingVariables	51 hospital services (emergency dept, postoperative recovery room, respiratory therapy, hospital auxiliary, physical therapy, blood bank, ICU, etc)	37 geriatric services (inpatient medical-surgical, rehab, psychiatric, swing beds, SNF, intermediate care facility, outpatient clinic, hospice, etc)	Output indicators (inpatient DRG weight, relative value scales for ambulatory as well as long-term care services)
Number of Hospitals	5,112 National Hospitals	416 hospitals /w elderly services	159 VA hospitals
Grouping Method	Hierarchical Cluster Analysis	Factor Analysis & Cluster Analysis	Ward's Cluster Analysis
Number of Peer Groups	17 groups using expected distinctiveness (dendrogram)	Seven groups (# hospitals ranged from 2 to 235)	Six groups (ranged from 14 to 45 hospitals per group) using R^2
Post-Evaluation	None	ANOVA /w exogenous variables	ANOVA /w exogenous variables
Application of Peer Groups	Compared with Barry's groups using rank correlation analyses	None	VA prospective payment system
Comments	Supported the hypothesis that there exists a well-defined hierarchy of hospital services	Geriatric services were represented by outreach, mental health, and inpatient factors	VA hospitals are budget-based, not profit-maximizers. Assumed a hybrid cost model

Appendix B: Peer Grouping Methods Used by PROs

The following table represents a rather incomplete record of peer grouping efforts as carried out by Peer Review Organizations (PROs). The information was received in part from direct contact with PRO statisticians, and in part from a request posted on the BIX by AMPRA.

PRO State	Peer Grouping Method(s)	Customized by project?	PRO Satisfaction	Provider Satisfaction
AR	Clustered using SAS (PROC CLUST) with inputs of bed size, case mix index (CMI), and services provided.	no		
	Were in the process of putting together peer groups based on cardiac services for the CCP, but the peer groups didn't get used.	yes		
CT	Used data from the POS file based on size and medical school affiliation to make a 2x2 table generating 4 peer groups.			
	Have a request into the AHA for more information, such as bed size, availability of CABG and/or catherization procedures.	yes		
Delmarva	8 peer groups based upon bed size and location of hospital. There were three sizes, under 100, over 400, and the rest. There were four locations: inner city, rural, suburban, and teaching. (This looks like there should be 12 groups, but evidently some cells were empty.)	no	high	
GA	Used peer groups as used by CT 'in the pilot based on cardiac services.' HCFA provided the peer grouping programs.	yes	high	high

PRO State	Peer Grouping Method(s)	Customized by project?	PRO Satisfaction	Provider Satisfaction
IN & KY	See Kwan Lee's paper. General idea is to form ranking of hospitals based primarily on size and CMI, split the list into 3 groups, and then split each group into two pieces based on factors measuring high tech/low tech.	no	high	medium high
	Broke up providers based on procedures performed at the provider in the preceding year. The procedures were ranked as follows: CABG≡Angioplasty≥Carotid Endarterectomy≥Pacemaker Insertion≥Hip Replacement or Laparoscopic Cholecystectomy Used solely for blood use project	yes	high	
KS	Three peer groups which are: small rural with basic services, medium regional referral centers, and large tertiary care centers.	no		
LA	Clustering based upon transfer rates into and out of the provider, and the service area population.	no	high	
	Tried to recompute the peer groups using current bed size, surgical admissions and medical admissions by using factor analysis	no		
MN	See attached (very detailed information). A brief summary: Tried some complicated and some simple methods, leaning towards using factor analysis to form scores by which the hospitals could be clustered.	no	lukewarm	lukewarm

PRO State	Peer Grouping Method(s)	Customized by project?	PRO Satisfaction	Provider Satisfaction
MS	Clustering based upon 4 variables: bed size, total number of hospital discharges, total number of Medicare discharges, and CMI.	no	high	
NY	Formed peer groups based on location (upstate/downstate), teaching status, size, as well as special designations of academic medical center, major public hospital, and exempt hospital. Formed groups in conjunction with individual hospitals and regional hospital associations.	no		
Pro-West	Use HCFA peer groups usually (size and urban/rural)			
	Have peer groups based upon number of discharges of the condition/procedure of interest.	yes		
TN	Clustered on bed size alone using the groups: 1-100, 101 to 300, 301 to 500, and 501+	no	high	high

Appendix C: Data Dictionary

AHA Data

Hospital Identification

AHA Unique Identification Number (ID)
Region Code, State Code

Organizational Structure

Control Code: Non-federal government
not-for-profit non-government
for-profit non-government
Federal government

Service Code: General medical and surgical, Pyschiatric, Others

Member System

Providers of Facilities and Services

General adult medical/surgical care
Medical/surgical intensive care
Cardiac intensive care
Burn care
Rehabilitation care
Alcohol/drug abuse or dependency
Psychiatric care
Intermediate care
Adult day care program
Alcohol/drug abuse or dependency outpatient services
Angioplasty
Arthritis treatment center
Assisted living
Breast cancer screening/mammograms
Cardiac catheterization laboratory
Case management
Community outreach
Crisis prevention

Providers of Facilities and Services (continued)

Dental services
Emergency Department
Certified trauma center
Extracorporeal shock-wave lithotripter
Fitness center
Freestanding outpatient center
Geriatric services
Health fair
Health information center
Health screenings
HIV-AIDS services
Home health services
Hospice
Hospital-based outpatient care center/services
Meals on wheels
Nutrition programs
Occupational health services
Oncology services
Open-heart surgery
Outpatient surgery
Patient education center
Patient representative services
Physical rehabilitation outpatient services
Primary care department
Psychiatric consultation/liaison services
Psychiatric education services
Psychiatric emergency services
Psychiatric geriatric services
Psychiatric outpatient services
Psychiatric partial hospitalization program
Radiation therapy
Computed-tomography scanner
Magnetic resonance imaging
Positron emission tomography
Single photon emission computed tomography
Ultrasound
Retirement housing
Social work services
Sports medicine
Support Groups
Transplant services

Providers of Facilities and Services (continued)

Transportation to health facilities
Urgent care center
Volunteer Services department
Women's health center/services

Participation in Physician Arrangement

Independent practice association
Group practice without walls
Open physician-hospital organization
Closed physician-hospital organization
Management service organization
Integrated salary model

Development of Insurance Products

Health maintenance organization
Preferred provider organization
Indemnity fee for service plan

Contracted Care (Blank if hospital did not respond to AHA Annual Survey)

Formal contract with HMO (HMO86)
Number of HMO contracts if HMO86
Formal contract with PPO (PPO86)
Number of PPO Contracts if PPO86

Total Facility Beds and Utilization

Total facility beds
Total facility admissions
Total facility inpatient days
Medicare inpatient discharges
Medicare inpatient days
Inpatient surgical operations
Outpatient surgical operations
Total surgical operations
Emergency outpatient visits
Other outpatient visits
Total outpatient visits

Inpatient Unit Beds (Setup and Staffed)

Number total hospital beds

Total Facility Expenses

Total payroll expenses
Employee benefits
Total hospital-unit payroll expenses
Hospital-unit employee benefits

Hospital Classification

Response Code
Community Hospital
Bed size code
Length of Stay Code

AHA Calculated Fields

Average daily census
Adjusted admissions
Adjusted patient days
Adjusted average daily census
Full-time equivalent physicians/dentists
Full-time equivalent registered nurses
Full-time equivalent licenses practical and vocational nurses
Full-time equivalent Medical and dental residents
Total Full-time equivalent hospital personnel
Total Full-time equivalent personnel

Geographic Codes -- Metropolitan statistical area code

Approvals

Accreditation by Joint Commission on Accreditation of Health Care Org
Cancer program approved by American College of Surgeons
Residency training approval by accreditation Council for GME
Medical School affiliation reported to American Medical Association
Member of Council of Teaching Hospitals of the Association of American Medical Colleges
Blue Cross contracting or participating
Medicare Certification by the Department of Health and Human Service

HCFA Data

Complexity of Services (Source: Case Mix Index File)

Case Mix Index

Medicare Patient Characteristics (Source: Medicare Part A Claims Data)

Number of discharges
Percent of female discharges
Percent of nonwhite race
Average patient age
Average length of stay
Percent transferred to other hospitals
Percent died in hospital
Percent transferred from other hospitals
Percent of emergency type admissions
Percent of one or more procedures
Female
Length of Stay
Died in Hospital
Emergency Admission: source
Emergency Admission: type
Caucasian
Transfers In: codes only
Transfers Out: codes only
Transfers In: dates and codes
Transfers Out: using dates
Had a Procedure Code
Case with valid age
Case with valid LOS
Case with known race
Case with known admission source
Case with known admission type
Count by mpn
Obs rate of female
Std. score of female, count: cases
Obs rate of died
Std. score of died, count: cases
Obs rate of emergency admission: source
Std. score of emergency admission: source, count: case with known admission source
Obs rate of emergency admission: type
Std. score of emergency admission: type, count: case with known admission type

Medicare Patient Characteristics (continued)

Obs rate of caucasians
Std. score of caucasians, count: case with known race
Obs rate of transfers in (codes only)
Std. score of transfers in (codes only), count: cases
Obs rate of transfers out (codes only)
Std. score of transfers out (codes only), count: cases
Obs rate of transfers out (dates and codes)
Std. score of transfers out (dates and codes), count cases
Obs rate of transfers out (dates and codes)
Std. score of transfers out (dates and codes), count: cases
Obs rate of transfers out (dates and codes)
Std. score of transfers out (dates and codes), count: cases
Obs rate of transfers out (dates and codes)
Std. score of transfers out (dates and codes), count: cases
Obs rate of having at least one procedure
Std. score of having at least one procedure procedure, count: cases
Average age at admission
Std Score of Age
Arithmetic Mean LOS
Std Score of ALOS
Geometric Mean LOS
Std Score of GLOS
Entropy

Physician Summary Information (Source: Physician UPIN File)

Number of physicians
Average years since medical school graduation
Apercent of active physicians
Percent of interns or residents
Percent of solo-practicing physicians
Percent of Medicare physicians
Average physician age
Percent of foreign medical school graduates
Percent of board-certified physicians

Appendix D

Face-validation of Peer Groups: Instructions to PROs

The statistical model for peer-grouping 5,326 acute care hospitals in the country was developed by using the administrative data that characterize the hospitals. The 20 variables which most discriminated the hospitals into clusters were (in order of their importance):

Length of stay	% Patients transferred out
Adjusted patient days	% Patients having a procedure
Number of Medicare discharges	MSA population size
Number of setup and staffed beds	Hospital occupancy rate
Total hospital FTEs	Accreditation by JCAHO
RN FTEs	Residents per average daily census
Average daily census	Membership in a hospital system
# of senior services provided	Participation in HMO
Case mix index	Participation in PPO
DRG entropy	For-profit ownership

Subjective validation of the clustering model is an important step in its development. The purpose of subjective validation is to ensure that the hospitals are indeed grouped according to their characteristics. This information is necessary for the utility of the model. It would suggest that the clustering methodology would be meaningful to the hospitals. Several PROs have volunteered to face-validate the peer-grouping of hospitals in their states. The statistical model has objectively grouped the hospitals, and the validators have to subjectively examine the appropriateness of the clustering (peer-grouping). The validator(s) should be the PRO personnel who is/are most familiar with the characteristics of the acute care hospitals in their state. The validation process is simply to read through the list of the hospitals and ensure that the hospitals with similar characteristics are grouped together and that the clusters somewhat differ from each other. At the state level, there may be a few, one, or none hospitals in some clusters. The validator(s) need not be concerned about it. Eventually, each state will determine their cut-off-points and collapse the 23 clusters into a much smaller number of peer-groups to suit the given situation. A second, 7-cluster model was developed by collapsing the 23 clusters. It is an example of a method providing less than 10 peer-groups. The face-validators are requested to validate both the 23 and 7 peer-groups of hospitals in their states.

The 23-cluster model was developed using the 1994 data on all acute care hospitals in the country participating in Medicare and providing a variety of services, general and/or specialized, and their Medicare patients. (In future, models will be developed to cluster hospitals providing specific services of interest.) As stated above, the statistical model accounted for all of the above 20 factors in clustering the hospitals. The validators are expected to use their personal knowledge of the hospitals in their respective states and validate the clustering models. Any disagreements with the clustering should be noted and the reasons for disagreements should be reported in detail using the attached standard forms.

Appendix D

Face Validation of Peer Groups: Instructions to PROs

State: _____

Hospital Name	Medicare Provider No.	Current Peer Group	Peer Group in Which Hospital to be Placed	Rationale for Change

Name of Reviewer: _____

E-mail Address: _____

Phone: _____

Fax: _____ Date: _____

Appendix E1: Seven Peer Groups (Clusters) by Statistics of Key Variables

Cluster	Statistics	Total Medicare Cases	Total Patient Days	Adjusted Patient Days	Set-up Staffed Beds	Average Daily Census	Occupancy Rate	Gerontology Services
A	Mean	549.1	2875.2	5001.2	68.56	37.10	0.46	2.98
	Std. Dev*	348.8	1971.8	3298.0	67.58m	59.48	0.21	2.16
	Min. Value	1.0	2.0	2.0	6.0	1.0	0.03	0.0
	Max. Value	1952	8095	15538	1193	1121	1.53	14.0
B	Mean	1697	10520	16477	140	76.57	0.54	6.46
	Std. Dev.	382.0	2237.1	3476.9	64.0	44.22	0.13	2.69
	Min. Value	693	4705	6314.8	40.0	27.0	0.18	0.0
	Max. Value	3146	17157	26215	493.0	380.0	1.0	16.0
C	Mean	2805	19154	28309	226.0	139.6	0.61	8.70
	Std. Dev.	594.3	3097.5	4806.7	123.1	106.1	0.13	3.05
	Min. Value	645	9058	9058	87.0	53.0	0.23	1.0
	Max. Value	4577	26583	59106	1983	1833	1.16	17.0
D	Mean	3892	27892	39969	294.3	188.8	0.65	10.5
	Std. Dev.	705.8	3005.4	4856.6	121.3	91.1	0.12	3.15
	Min. Value	1746	15371	28513	102	84.0	0.30	0.0
	Max. Value	6266	35430	69298	1150	893.0	1.22	18.0
E	Mean	5257	39755	55045	370.7	250.9	0.69	11.5
	Std. Dev.	983.0	5113.8	7496	140.6	98.96	0.11	2.91
	Min. Value	2016	25392	36693	171.0	67.0	0.18	3.00
	Max. Value	8450	54085	97300	1446	1043	1.04	18.0
F	Mean	7559	61972	81812	539.2	393.8	0.74	13.0
	Std. Dev.	1647	10634	13861	171.9	131.0	0.10	2.82
	Min. Value	3428	42062	59017	247.0	197.0	0.42	6.0
	Max. Value	12376	92663	117644	1480	962.0	1.13	18.0
G	Mean	121186	106195	137826	847.5	639.8	0.76	14.7
	Std. Dev.	2199	16836	24035	246.1	228.3	0.17	2.49
	Min. Value	8449	85553	109955	350.0	127.0	0.16	7.0
	Max. Value	19268	154940	219376	1407	1203	1.27	18.0

* Standard Deviation

Appendix E2. Twenty-three Cluster Model by Statistics for Key Variables

Cluster	Statistics	Total Medicare Cases	Total Patient Days	Adjusted Patient Day	Set-up Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
A	Mean	262.55	1223.2	2280.6	50.75	27.49	0.44	1.83
	Std. Dev.*	145.88	720.92	1377.7	41.58	37.17	0.24	1.63
	Min. Value	1.0	2.0	2.0	6.0	1.0	0.03	0.0
	Max. Value	715.0	3661.0	8220.6	611.0	480.0	1.53	14.0
B	Mean	834.50	4505.9	7705.3	89.86	49.56	0.50	4.21
	Std. Dev.	182.92	819.62	1658.5	93.04	85.53	0.17	1.95
	Min. Value	49.0	402.0	402.0	22.0	6.0	0.15	0.0
	Max. Value	1952.0	6325.0	15538	1193	1121	1.37	11.0
C	Mean	1402.8	8438.4	13453	120.7	63.62	0.53	5.77
	Std. Dev.	231.16	1045.5	1865.1	54.72	38.52	0.14	2.41
	Min. Value	693.0	4705.0	6314.8	40.0	27.0	0.18	0.0
	Max. Value	2034.0	11198	19766	444.0	299.0	1.0	14.0
D	Mean	596.20	3041.2	5419.9	65.60	34.90	0.48	3.32
	Std. Dev.	117.45	556.46	1174.7	45.80	42.63	0.20	1.79
	Min. Value	96.0	1162.0	1162.0	14.0	9.0	0.17	0.0
	Max. Value	916.0	5572.0	9879.1	460.0	474.0	1.29	9.0
E	Mean	4229.3	30567	43452	326.4	211.7	0.65	10.97
	Std. Dev.	654.55	2093.0	3955.4	148.7	111.9	0.10	3.04
	Min. Value	1767.0	23321	34427	150.0	99.0	0.30	1.0
	Max. Value	6266.0	34831	54445	1150	893.0	0.93	18.0
F	Mean	2795.9	18960	27731	222.7	138.9	0.62	8.61
	Std. Dev.	493.10	1822.7	2912.4	101.1	83.44	0.13	2.80
	Min. Value	1291.0	10896	18812	91.0	66.0	0.23	2.0
	Max. Value	4023.0	24486	38586	1075	965.0	1.16	16.0
G	Mean	3715.1	26488	38140	277.4	176.7	0.64	10.19
	Std. Dev.	667.36	2398.1	4252.6	100.3	75.53	0.12	3.18
	Min. Value	1746.0	15371	28513	102.0	84.0	0.30	0.0
	Max. Value	5390.0	35430	69299	620.0	523.0	1.22	17.0

Cluster	Statistics	Total Medicare Cases	Total Patient Days	Adjusted Patient Day	Set-up Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
H	Mean	1706.5	10468	16544	144.7	79.17	0.54	6.52
	Std. Dev.	237.75	997.47	2034.7	66.09	48.41	0.12	2.71
	Min. Value	888.0	7632.0	9314.0	57.0	30.0	0.28	1.0
	Max. Value	2291.0	13723	21355	493.0	380.0	0.89	15.0
I	Mean	1114.5	6316.6	10422	105.9	56.47	0.50	4.88
	Std. Dev.	205.98	910.54	1730.8	93.72	83.30	0.16	2.15
	Min. Value	109.0	479.0	622.68	30.0	19.0	0.22	0.0
	Max. Value	1929.0	8095.0	14342	1093	1086	1.27	13.0
J	Mean	7298.4	58454	78099	532.4	381.4	0.72	13.07
	Std. Dev.	1245.5	4228.1	7278.20	147.6	116.5	0.11	2.95
	Min. Value	4233.0	48017	60279	247.0	226.0	0.42	6.0
	Max. Value	9893.0	70541	108530	1273	962.0	1.13	18.0
K	Mean	2358.5	15836	23973	197.3	119.2	0.59	7.79
	Std. Dev.	406.90	1316.5	2574.8	148.9	138.2	0.13	2.84
	Min. Value	645.0	9058.0	9058.0	88.0	57.0	0.32	1.0
	Max. Value	3655.0	19452	29948	1983	1833	1.06	15.0
L	Mean	3215.3	22333	32826	255.3	158.4	0.62	9.61
	Std. Dev.	546.81	1909.5	4079.6	114.0	91.93	0.13	3.24
	Min. Value	1475.0	11627	17919	87.0	53.0	0.26	3.0
	Max. Value	4577.0	26583	59106	1236	993.0	0.98	17.0
M	Mean	2031.3	12971	19935	160.6	89.75	0.57	7.21
	Std. Dev.	323.32	1159.1	2267.6	65.84	43.22	0.12	2.79
	Min. Value	972.0	9565.0	14480	67.0	42.0	0.20	2.0
	Max. Value	3146.0	17157	26215	462.0	300.0	0.87	16.0
N	Mean	5377.2	39725	55002	373.7	256.6	0.69	11.57
	Std. Dev.	858.12	2302.4	5098.1	163.4	116.5	0.09	2.69
	Min. Value	2016.0	30608	36693	171.0	142.0	0.48	4.0
	Max. Value	7250.0	46912	69520	1446	1043	1.01	16.0

Cluster	Statistics	Total Medicare Cases	Total Patient Days	Adjusted Patient Day	Set-up Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
O	Mean	4735.0	35260	49392	351.2	232.5	0.67	11.35
	Std. Dev.	796.61	2387.7	4726.7	132.5	97.49	0.11	3.03
	Min. Value	2416.0	29716	38762	189.0	67.0	0.18	4.0
	Max. Value	6934.0	42506	62132	1150	940.0	1.04	18.0
P	Mean	5779.8	45057	61715	391.4	268.3	0.70	11.68
	Std. Dev.	966.65	3734.7	6065.2	128.7	82.02	0.12	2.92
	Min. Value	2903.0	25392	52249	202.0	123.0	0.22	3.0
	Max. Value	8450.0	54085	97300	965.0	697.0	1.02	17.0
Q	Mean	6522.0	52077	68724	449.0	323.4	0.73	12.0
	Std. Dev.	1171.1	3021.0	5070.6	132.9	103.5	0.09	3.0
	Min. Value	3428.0	42062	59018	251.0	197.0	0.52	6.0
	Max. Value	8978.0	59181	83090	881.0	831.0	0.94	18.0
R	Mean	8021.1	68940	90891	574.6	423.0	0.74	13.39
	Std. Dev.	1389.5	4751.4	6776.5	145.0	105.6	0.09	2.43
	Min. Value	4663.0	59356	78883	357.0	279.0	0.50	7.0
	Max. Value	11695	78148	117321	978.0	794.0	0.94	17.0
S	Mean	11259	92808	119126	723.8	546.1	0.78	14.65
	Std. Dev.	1324.1	5069.1	6082.6	208.8	126.0	0.17	2.76
	Min. Value	8739.0	85553	109955	350.0	236.0	0.60	7.0
	Max. Value	13865	101114	132504	1088	762.0	1.27	18.0
T	Mean	9682.9	80826	104658	692.7	528.4	0.78	14.33
	Std. Dev.	1569.9	6146.1	8130.3	215.9	140.7	0.10	1.91
	Min. Value	6964.0	69381	88722	314.0	279.0	0.55	10.0
	Max. Value	12376	92663	117644	1480	852.0	0.95	18.0
U	Mean	12024	103525	134781	798.6	610.4	0.77	14.71
	Std. Dev.	2229.1	3856.2	9314.7	136.2	170.8	0.16	1.44
	Min. Value	8449.0	97482	123560	560.0	290.0	0.34	13.0
	Max. Value	14649	110587	151334	1032	932.0	0.95	18.0

Cluster	Statistics	Total Medicare Cases	Total Patient Days	Adjusted Patient Day	Set-up Staffed Beds	Avg. Daily Census	Occupancy Rate	Gerontology Services
V	Mean	13468	120623	157204	1017	735.2	0.71	14.56
	Std. Dev.	2769.6	4504.6	9468.8	245.9	320.1	0.24	3.61
	Min. Value	10456	114542	142142	596.0	127.0	0.16	7.0
	Max. Value	19268	130164	168831	1394	1134	0.95	18.0
W	Mean	14349	151237	199864	1270	1022	0.81	15.0
	Std. Dev.	1616.8	4021.5	20325	133.2	164.3	0.09	1.73
	Min. Value	12972	146959	178813	1141	882.0	0.70	14.0
	Max. Value	16129	154940	219376	1407	1203	0.86	17.0

* Standard Deviation

Appendix F: User's Guide

1. Conventions Used in This Manual

The users of the Health Care Financing Administration's (HCFA's) Standard Data Processing System (SDPS) will initially be the users of this model. SDPS is a rigidly controlled system in which all users at all sites have (nearly) the same systems, with variations occurring only by groups of users. All SDPS users are using Microsoft Windows 3.11 and SAS 6.11 for Windows. Consequently, references to specific drive letters and directories are given below. If SDPS is not used, but MS-Windows or MS-Windows95 is used, the references to specific drive letters and paths may be ignored and replaced with those more suitable for the system. If a Macintosh or OS/2 is used, the word 'folder' should be substitutedfor each instance of the word 'directory' below. If a Macintosh is used, the word 'volume' should be substituted for the word 'disk', and the 'name of the drive' be substituted for the 'drive letter'. If some variety of Unix is being used, all references to drive letters should be ignored. Lastly, your operating system's file separator should be used wherever the '\' character appears in a path. For the Macintosh, this is ':', and for Unix, this is '/', and for OS/2 it is '\'. The following instructions are based on the assumption that the user of this model is using the SDPS system. As for typography, computer file names, input and output will be specified by a `fixed-width font`. Generic computer input for which the user must replace text will be in *`an italicized fixed-width font`*. Important items will be **bold**. Items which are still in limbo will be denoted by *italics*. Lastly, items which may not appear in the final version will be [included between brackets].

2. Overview of the Software

2.1. Purpose of the Software

This software uses data concerning individual hospitals to form groups of similar hospitals called hospital peer groups. While there are any number of ways to accomplish this task, the software is designed with several goals in mind. First of all, the software should allow the user to easily use sophisticated statistical clustering techniques. Consequently, programming by individual users may not be necessary. Next, the software is designed to allow for easy experimentation. This should allow the curious to find techniques which will work for them, and could work well for all. Lastly, by distributing the software freely to users, it will allow more portability of projects which requires special peer grouping.

2.2. What is Needed?

2.2.1. Data Requirements

There must be a hospital-level data set, i.e., a data set which contains characteristics for each hospital which will be clustered. There are no requirements for either the number or type of hospitals included in the data set. There are no requirements for the types of characteristics included in the data set, except that all characteristics must be numerically coded. Since the

analyses are written to use the SAS statistical package, the data must be in a format which can be understood by SAS.

2.2.2. Time Requirements

There are two major time requirements -- those of the user and those of the computer. Once the data set is in its proper form, the user can expect to spend as little as five hours or as many as five thousand hours to analyze the data;[1] a reasonable budget would include *40 hours* of an analyst's time. The computer time depends heavily on the type of processor and platform. [For careful analysis, including exhaustive bootstrapping, the user can expect to let the computer run between 25 and 35 hours.] This package requires about 1.5 hours to read the manual and install the software, and 0.5 hours for each run.

2.2.3. Required Knowledge

The user should be familiar with both factor analysis and cluster analysis, though he/she need not be an expert in either technique. Since the program is written to run within the SAS statistical package, some knowledge of SAS is assumed, though this, too, need not be extensive. Lastly, the user needs either an intimate knowledge of a reasonable subset of the hospitals being clustered or access to a person with this knowledge who would be willing to take a subjective look at the final peer groups.

2.3. Questions to Ask Before Using

2.3.1. What Does This Software Do?

This software forms hospital peer groups by first reducing the hospital-level characteristics to factor scores through the use of factor analysis, and then by clustering the hospitals into peer groups using cluster analysis. There is an additional element to the program which helps choose the proper number of peer groups by using the bootstrapping technique (Jain and Moreau, 1987; Reference 23 in the Report). This software requires the user to explicitly use his/her judgement in *three* places during the processing.

2.3.2. What Does This Software Not Do?

Though there are default values for all of the settings which may be tweaked by the user, blind use of these defaults will most likely not lead to meaningful clusters.

2.3.3. Instructions for the Begining SAS User

The beginning SAS user will need to read about `proc factor` and `proc cluster`, so that he/she can learn about the available options and output from these procedures. The beginner needs to know how to start SAS, submit jobs, and read SAS output.

3. Getting Started

---

[1]The upper limit is, just as for all statistical analysis and investigation, limited only by the time available to the analyst.

3.1. Organizing Your Data

3.1.1. What Data are Needed?

All that is needed is a data set which contains numerical characteristics of hospitals. The data set should contain as few missing values as possible, since missing values will lead to the exclusion of those hospitals from being clustered.

3.1.2. What Data Structure is Required?

There is no requirement for the number of variables needed. The program is simple enough that it can handle whatever variables are wanted. The usual statistical caveats about parsimony do apply! If the data are already at the hospital level, then SAS will use approximately 3.5 times the space of the original data set. This additional space is used for the ubiquitous temporary files containing copies of the original dataset and output from the various procedures, and which are in use for a SAS session. These files will disappear when SAS is exited normally.

3.2. Organizing the Computer

A directory to hold all your work should be created. Its name is immaterial, though it needs to be noted for setting up the program properly. No particular structure is needed, though it is strongly suggested that you use care in saving the temporary data sets to permanent data sets, because SAS doesn't protect against overwriting files.

4. Doing the Work

4.1. Preparing the Program

Very little needs to be done to the program other than installation. Since the program is written as a SAS macro, it is flexible and, in most instances, will not need to be edited by the user. Editing, the code is not precluded.

4.1.1. Installing the Program in the Proper Place

Since this is a SAS macro which will be called from other SAS programs, it needs to installed in a place where SAS can find it. This requires two things to be done. First, a central collection point for user-defined macros needs to be made. This collection point must **not** be in the SAS application directory, and the `config.sas` file needs to be edited for each user. Once again, the instructions below are intended for users of SDPS. If SDPS is **not** used, all references to particular drive letters and directories should be ignored. Instead, a place for all SAS customizations should be chosen, and those drives and paths should be substituted for the ones given below.

4.1.1.1. Create a Central Directory for SAS Macro

To begin installation, a directory to hold SAS macros should be created. All users should have the directory available to them and should be able to write to it. One way to do this is to open

File Manager, and then open a window for the s: drive. Click once with mouse button 1[2] on the sasdata folder icon in the left frame. Then click on the File menu, and select the Create Directory... item. When prompted for the directory name, put in whatever name you'd like that fits into DOS' 8 character limit. By using the s: drive, all SAS users will have access to the same copy of the file, assuring synchonization. For the purposes of having a concrete name in this manual, usrmacro (for User's Macros) will be used for demonstration purposes.

4.1.1.2. Copy the Programs to Directory

A copy of all the files from the programs directory should be put into the newly created directory. The names of any of the files should not be changed, since this will prevent SAS from finding and running them.

4.1.1.3. Let SAS Know Where Files Are

After the usrmacro directory has been set up, it should be ensured that SAS knows of the directory's location. To do this, File Manager should be used to find the place where SAS is installed on the system. Typically (and unfortunately), it is often at C:\SAS). A copy of the config.sas file should be made to protect against any disasters. Once that is done, the config.sas with a text editor such as Notepad or the SAS editor should be opened. A word processor should **not** be used. If SAS is not already running, double-clicking the icon will start SAS, once the config.sas is in the program editor. If SAS is already running, double-clicking the icon will give an error message. The following line should be inserted at the top of the file:

```
-SET common s:\sasdata
```

where s:\sasdata\usrmacro is the **full** path to, but does not include the directory made in Section 4.1.1.1. Next, the text sasautos should be searched twice. The second search should find a line which looks like: `-SET SASAUTOS(`

Scroll down to the closing parenthesis, and insert a line before the closing parenthesis which reads:

```
!common\usrmacro
```

where you insert the name of the directory from Section 4.1.1.1 in place of usrmacro. The file should be saved. If you were in SAS, exit.

Start SAS and type `%customx;` in the program editor. Submit the line. The log file should display a line telling that everything is set up properly. If desired, open the autoexec.sas file which is in the same directory as the config.sas file. Add the line `%customx;` at the bottom. In this way, you'll always know if something has happened which would cause you to lose access to the added macros.

The second alteration to the config.sas file occurred below the line exhorting the user not to edit anything. This editing is safe, but may be destroyed if any additional installations to SAS are done.

4.1.2. Editing the Macro

---

[2]Mouse button 1 is what is often called the 'left mouse button.' This is a misnomer, since it is the right mouse button for those of us who use the mouse with our left hands. Either way, it is the 'main mouse button.'

The macro can be edited. However, be sure that your editing does not affect the original or affect the behavior of the macro for others. When you edit the macro, change its name (the word after `%Macro` in the first line), and then save the file so that the filename matches this new name. For example, if you change the name of the macro to `foo` then be sure to save the file as `foo.sas`.

4.2. Using the Program

4.2.1. General Usage Rules

Since this program is coded as a macro, it can be used either alone or embedded in a larger SAS program. It can be used to mimic the work of the peer grouping workgroup, or can have its factor analysis options and cluster analysis options changed.

The basic usage of the `peer6` macro is similar to that of other macros in SAS. It can be called by inserting a line

```
%peer6(parameters...);
```

in the editor window (or in a program file). The parameters must be specified according to the following rules. As written, there are 14 parameters which can be set, most of which have default values and can be omitted entirely. These parameters with examples are explained below.

Parameter Name	**Usage**	**Default Value**
`indir`	Directory in which the hospital-level SAS data file sits. The path must be enclosed in single quotes. If this is blank, then the macro **must** be called from within a data step. If non-blank, then the `file` option must also be used.	blank
`file`	This is the name of the hospital-level data file to be used. If this is blank, then the macro **must** be called from within a data step. If non-blank, then the `indir` option must also be used.	blank
`mimic`	If non-empty, the action of the macro will be to mimic the peer grouping used by the members of the peer grouping workgroup. For the option to work correctly, a dataset with the same variables as that used by the workgroup is needed. The option is not picky about its contents; anything[3] other than a blank will cause this option to execute.	blank

[3]Including the name of Aunt Mergatroid's pet cat....

Parameter Name	Usage	Default Value
varlist	The variables to be used for the factor analysis. **This is required** unless mimic is used. If omitted, SAS will continue to work creating several pages of worthless output.	blank
facopts	Whatever options are desired for the factor analysis; these are passed on to proc factor, whose usage can be found in the SAS manual. The score option **must** be specified if any facopts are given.	method=ml maxiter=1000 msa rotate=promax reorder score flag=.6
factors	The number of factors to keep.	1
method	The clustering method to use.	ward
clusopts	Options for whatever clustering method is chosen.	std pseudo ccc print=30
clusters	The number of clusters to keep.	23
mpnmin	Smallest Medicare Provider Number (MPN) to print when printing a subset. Must be specified together with mpnmax. For subsetting to work properly, there must be three variables present in the main dataset: mpn, which contains the Medicare provider number, mname which contains the provider name, and mloccity which contains the provider city.	blank
mpnmax	Largest (MPN) to print when printing a subset. Must be specified together with mpnmin.	blank
outdir	Directory to store datasets generated within the program. The full path to the target directory must be specified within single quotes. If this is not specified, then the datasets will be stored in the SAS Work library, and hence will be temporary.	blank
outlib	A name for the output library. Use this with the outdir option to name the library in the SAS fashion.	outlib
sasname	The name to be used for the main data file within SAS. This can be used to avoid naming conflicts when the macro is used within a larger SAS program. The name given to sasfile determines the name of the main data set. If dir and file aren't used, then either this must be specified, or the working file must be named Peer.	Peer

Parameters are specified by *parameter=value* and are separated by commas. Lists of values for a

single parameter, such as a list of variable names given to `varlist` or options given to `clusopts`, can be separated by spaces or newlines. The order in which the parameters are specified is unimportant, since all the parameters have names.

As a final usage note, probably the best use of the macro would be to have three pieces included in a program which used this macro: (1) a data step which read in your data set and manipulated the data, (2) the `peer6` macro, and (3) a final step which played with the results.

4.2.2. Examples

4.2.2.1. Mimicking the Workgroup Peer Grouping

The following should be used to mimic the analysis done by the workgroup on the SAS file `hosdata.sd2` which sits in the directory `d:\current\peergrp\data`.

```
%peer6(`dir=d:\current\peergrp\data',file=hosdata,mimic=foo);
```

Since the analysis is the same as that of the workgroup, the variables are also taken as the same as those of the workgroup.

4.2.2.2. Use of Macro Within a SAS Program

To embed the macro inside a larger SAS program, the following is the minimum needed to (once again) mimic the actions of the workgroup:

```
LIBNAME MainLib d:\current\ropeer\data;
DATA Main;
SET MainLib.mrg1;
%peer6(mimic=bar);
```

This has exactly the same effect as the first example. Once again, the `varlist` is not needed because of the `mimic` option.

4.2.2.3. Use of Different Variables and Options

A customized call to the macro which uses a different set of variables and keeps 2 factors in the factor analysis would be

```
%peer6(`dir=d:\foo\bar',
file=afile,
factors=2,
varlist=bedsize avgdcen
mlos transout transin entropy,
);
```

This runs the analysis on the dataset `d:\foo\bar\afile.sd2` using 2 factors generated by the variable list given above.

4.3. Results

4.3.1. Data Sets

Running the macro generates *5* data sets in the working directory. The names and the contents follow.

File	Contents
Peer[4]	The data file on which the analyses are run.
FacRes	The results of the factor analysis.
Scoring	The combination of the main data set with the factor scores for each of the providers. The variables holding the factor scores are named factor1, factor2, etc.
Tree	The information from the clustering which would be needed to generate the dendogram for the clustering.
TreeRes	The results of running proc Tree on the Tree dataset.
ScoTree	The final main file, which has the original data plus the factor scores plus the cluster numbers and symbols for each observation. This is the dataset which would be used in further analysis.

Note that unless the outdir option is used, all of the above files are temporary files (i.e., in the work library), and hence will disappear when SAS is exited. The filenames are purposely hardcoded.

4.3.2. Printed Output

The output consists of values for the options, followed by output from the factor analysis, followed by output from the cluster analysis, followed by means of the variables used in the factor analysis. This output appears for each peer group. If the mimic option is used, there are also some text plots of various variables by the factor scores.

5. Experimenting

As noted above, the datasets which result from the factor and cluster analysis are kept in temporary files by default. If you wish to save the resulting data sets, specify a directory for the outdir parameter. SAS does not care if it overwrites existing files, so if you specify a directory which already contains files which you want to keep, the earlier files will be overwritten. The only precaution you can take is to be very careful!

6. Troubleshooting

Symptom: peer6 won't run, and instead comes back with an error

```
WARNING: Apparent invocation of macro PEER6 not resolved.
  (a listing of the offending line)
ERROR 180-322: Statement is not valid or it is used out of proper order.
```

4

Disease: This error occurs when SAS cannot find the peer6 macro.
Cure: There are a few possibilities for how this could happen. First check that the peer6 macro was copied to an accessible directory. Then check to be sure that the config.sas file was altered so that it could find the directory. See Section 4.1.1.1 on page 3 for details.

Symptom: peer6 complains of a missing module, and the error
```
ERROR: Module PEER      not found in search paths.
ERROR: Error in the LIBNAME or FILENAME statement.
```
The error is located near the top of the log file, and where *PEER* is the sasname you specified or PEER if no sasname was specified.
Disease: SAS is becoming confused when the libname *PEER* is being set to point to a directory. This is caused by missing quotes around the indir directory.
Cure: Put single quotes around the directory specified by the indir parameter.

Symptom: peer6 complains of a missing module, and the error
```
ERROR: Module OUTLIB      not found in search paths.
ERROR: Error in the LIBNAME or FILENAME statement.
```
The error is located near the top of the log file, and where *OUTLIB* is the outlib you specified or OUTLIB if no outlib was specified.
Disease: SAS is becoming confused when the libname *OUTLIB* is being set to point to a directory. This is caused by missing quotes around the outdir directory.
Cure: Put single quotes around the directory specified by the outdir parameter.

Symptom: peer 6 complains that a data set is missing, giving an error such as
```
ERROR: File PEER.FOOBAR.DATA does not exist.
```
Disease: The file which was to be used for input, and which was specified as the file parameter does not exist in the directory where you told SAS it would be found.
Cure: Fix either the indir or the file parameter so that the correct file is being used.

Symptom: the macro doesn't run, but instead give an error
```
ERROR: The keyword parameter SOMETHING was not defined with the macro.
```
Disease: A nonexistent or misspelled parameter was used.
Cure: Check the spelling in the call to the macro. Check also that commas occur only between values given to parameters, and not within the lists of values given to a single parameter.

Some other common typing errors which create problems with SAS are:

- Using semi-colons to separate parameters rather than commas will cause unpredictable behavior.
- Leaving one of the two quotes needed around a path will cause the macro to sit patiently and wait.
- Leaving out a comma between two parameters will cause an unpredictable mess.

CMS LIBRARY
3 8095 00013273 4

U.S. Department of Health and Human Services
Health Care Financing Administration
Baltimore, Maryland
March 1999
HCFA Pub. No. 10130